Optimization-based remap and transport: a divide and conquer strategy for feature-preserving discretizations

Pavel Bochev[1]

Numerical Analysis and Applications, Sandia National Laboratories, MS-1320, Albuquerque, NM 87185-1320, USA

Denis Ridzal[1,*]

Optimization and Uncertainty Quantification, Sandia National Laboratories, MS-1320, Albuquerque, NM 87185-1320, USA

Kara Peterson[1]

Numerical Analysis and Applications, Sandia National Laboratories, MS-1320, Albuquerque, NM 87185-1320, USA

Abstract

This paper examines the application of optimization and control ideas to the formulation of *feature-preserving* numerical methods, with particular emphasis on the conservative and bound-preserving remap (constrained interpolation) and transport (advection) of a single scalar quantity. We present a general optimization framework for the preservation of physical properties and relate the recently introduced flux-variable flux-target (FVFT) [1] and mass-variable mass-target (MVMT) [2] optimization-based remap (OBR) to this framework. Both cast remap as a quadratic program whose optimal solution minimizes the distance to a suitable *target* quantity, subject to a system of linear inequality constraints. An approximation of an exact mass update operator defines the target quantity, which provides the best possible accuracy of the new masses without regard to any physical constraints such as conservation of mass or local bounds. The latter are enforced by the system of linear inequalities. In so doing, OBR separates accuracy considerations from the enforcement of the physical properties. We follow with a formal examination of the relationship between the FVFT and MVMT formulations. Using an intermediate flux-variable mass-target (FVMT) formulation we show the equivalence of their optimal solutions.

To underscore the scope and the versatility of the OBR approach we introduce the notion of adaptable targets, i.e., target quantities that reflect local solution properties, extend FVFT and MVMT to remap on the sphere, and use OBR to formulate adaptable, conservative and bound-preserving optimization-based transport algorithms, both in $\mathbb{R}^d$ and on the sphere. A selection of representative numerical examples demonstrates the computational properties of our approach.

Keywords: constrained interpolation, remap, quadratic program, preservation of local bounds, preservation of linearity, optimization-based remap, transport, spherical coordinates.

---

[*] Corresponding author

Email addresses: `pbboche@sandia.gov` (Pavel Bochev), `dridzal@sandia.gov` (Denis Ridzal), `kjpeter@sandia.gov` (Kara Peterson)

URL: `http://www.sandia.gov/~pbboche/` (Pavel Bochev), `http://www.sandia.gov/~dridzal/` (Denis Ridzal)

[1] Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Preprint submitted to Journal of Computational Physics *August 9, 2012*

1. Introduction

Fundamental physical properties of natural phenomena give rise to the salient analytical properties of their mathematical models. However, the discretization of these models can cause the loss of key mathematical relationships, potentially leading to ill-posed discrete equations, the emergence of spurious modes, or physically impossible solutions. Coupled multiphysics simulations, where the output from one constituent component provides the input to another component, can further exacerbate the loss of structural and qualitative information in the discrete model. There, an unphysical solution from one component can cause a cascading breakdown of the subsequent components and compromise the whole simulation.

Thanks to extensive research efforts over the past 20 years it is now well understood that a discrete "vector calculus" setting, i.e., discrete spaces and operators that mimic basic vector calculus properties such as the Poincare lemma, the Gauss divergence theorem and the Stokes circulation theorem, enables *structure-preserving* discretizations of a large class of Partial Differential Equations (PDEs); see [3, 4, 5, 6, 7] and the references therein.

However, advances in structure-preserving numerical methods stand in sharp contrast with the limited mathematical and algorithmic understanding of *feature-preserving* discretizations, i.e., discretizations that reproduce qualitative properties of the exact solutions, such as local bounds, maximum principles, and symmetries, to name a few. The root cause for this imbalance is that such properties emerge from the interplay of function space structures with differential and boundary operators, i.e., they are model specific, whereas the given functional space structures can be shared by multiple mathematical models. For instance, the same discrete vector calculus setting can support the structure-preserving discretization of models ranging from pure diffusion to conservation laws, yet the unique qualitative properties of the models at each extreme, such as the maximum principle and local bounds are not guaranteed to emerge automatically from that setting.

The fluid nature of the qualitative properties makes it difficult to enforce them directly in the discretization process, that is by relying solely on the mesh structure and the discrete variables. As a rule, this strategy ties together the preservation of the desired features with geometric conditions on the mesh and/or restrictions on the accuracy. A typical example is the discrete maximum principle (DMP) for the Poison equation, which requires a monotone, or M-stiffness matrix. To ensure this property on triangular elements the sum of the two angles opposing each interior edge should be less than π, and the polynomial degree should be 1 [8, 9, 10]. An extension of DMP to more general triangular or quadrilateral meshes requires nonlinear modifications of the governing equations such as the nonlinear stabilized finite element method for the Poisson equation [11, 12], the nonlinear extension of the diamond scheme [13], and the nonlinear finite volume scheme in [14].

High-order maximum-principle satisfying and positivity preserving schemes for conservation laws exist in one dimension [15], or on rectangular meshes [16, 17]. An extension of these schemes even to triangular elements is highly nontrivial [18]. A similar interdependence between mesh, accuracy and preservation of a physical property exists in many of the slope and flux limiters in use today. As a result, many of them do not preserve linear functions on irregular meshes [19], which impacts accuracy and robustness. This interdependence is propagated to any algorithm that employs limiters such as advection-based remappers in Arbitrary Lagrangian Eulerian (ALE) methods [20].

This paper draws upon and continues our previous efforts to develop an alternative, optimization-based *divide-and-conquer* strategy [21, 22, 23, 24] for the formulation of stable, accurate and physically consistent discretizations. Specifically, here we focus on the application of optimization and control ideas to separate stability and accuracy considerations from the enforcement of the desired physical properties. In a nutshell, given a mathematical model and a list of desirable physical properties, our approach seeks the corresponding discrete model in the form of a constrained optimization problem in which

- the objective is to minimize the distance, measured in some suitable norm, between the discrete solution and a given *target* solution;
- a discrete model that is *stable and accurate* but is not expected to posses all desired physical properties defines the target solution;

- the optimization constraints *enforce* any desired physical properties that are not already present in the target solution.

This strategy offers a number of important theoretical and computational advantages in the formulation of feature-preserving numerical methods:

- the numerical solution is a global optimal solution from a feasible set defined by the desired physical properties, i.e., it is always the best possible, with respect to the target, approximate solution that also possesses these physical properties;
- the decoupling of the target definition from the preservation of the physical properties allows to adapt the numerical solution to different problems by choosing the most appropriate target definition and objective function for these problems;
- the enforcement of the desired properties as optimization constraints is impervious to the mesh structure and/or field representations, thereby enabling feature-preserving methods on arbitrary unstructured meshes, including polygonal and polyhedral meshes.

The present work applies the optimization-based strategy to the high-order accurate and feature-preserving remap (constrained interpolation) and transport (advection) of a single scalar conserved quantity ("mass"). The features that we aim to preserve through the use of optimization are (a) the conservation of total mass and (b) physically motivated local bounds on the primitive variable (the density). The remap task arises in Arbitrary Lagrangian-Eulerian (ALE) methods, where high-order remapping between meshes is critical for the accuracy of the simulation, especially in conjunction with a *continuous rezone* approach, which requires remapping at every time step [25, 26]. The second task, i.e., the stable, accurate and feature-preserving solution of transport equations on general unstructured meshes, remains one of the most challenging numerical problems. Our work exploits the intrinsic connection between transport and incremental remapping [27, 28, 29] to extend the optimization-based remap to conservative, linearity and local bounds preserving transport algorithms.

It should be noted that optimization ideas also motivate the Flux-Corrected Remap (FCR) [30] although FCR itself is algorithmically not an optimization-based remap. Nonetheless, the ideas outlined in [30] prompted the development of bona fide optimization-based formulations of remap [1, 24, 2]. The paper [1] also examines the relationship between such formulations and FCR, showing that FCR admits interpretation as an approximate *solution algorithm* for an optimization-based remap problem. Elements of an optimization-based approach can be found in one form or another in some recent work on the optimization-based preservation of DMP [31, 32, 33, 34].

The paper starts with a brief summary of the basic ideas of the optimization-based approach for feature-preserving discretizations in Section 2. Section 3 introduces the relevant notation and background results, and defines the remap problem we seek to solve. Section 4 specializes the abstract framework into a generic optimization-based remap (OBR) formulation and then examines two instances of OBR corresponding to the recent *flux-variable flux-target* (FVFT) [1], and *mass-variable mass-target* (MVMT) [2] methods. We also introduce an intermediate *flux-variable mass-target* (FVMT) formulation, which bridges FVFT and MVMT. Section 5 studies the theoretical properties of the OBR formulations, whereas Section 6 demonstrates the versatility of OBR by introducing the notion of *adaptable targets*, and specializing OBR to spherical geometry and extending it to feature-preserving, optimization-based transport (OBT). Section 7 contains representative numerical results that illustrate the accuracy, efficiency and versatility of the optimization-based approach. Section 8 summarizes our conclusions.

2. Optimization-based approach for feature-preserving discretizations

This section briefly discusses an abstract optimization-based framework for feature-preserving discretizations. The setting follows and specializes the ideas explored in [22, 23, 1]. We study the approximate numerical solution of the abstract operator equation: find $u \in X$ such that

$$L(u) = f \quad \text{in } Y, \tag{2.1}$$

where X and Y are two Banach spaces, $L : X \mapsto Y$ is a given operator, and $f \in Y$ is the given data. The main objective is to preserve the properties of the exact solution and the operator that are deemed essential for predictive simulations. We make the following assumptions about our abstract problem:

A.1 There exist subsets $\mathscr{U} \subseteq X$ and $\mathscr{V} \subseteq Y$ such that for any $f \in \mathscr{V}$ the solution of (2.1) satisfies $u \in \mathscr{U}$.

A.2 There are discrete spaces X_h and Y_h, parametrized by $h > 0$, and a sequence of discrete operators $L_h : X_h \mapsto Y_h$ such that the discrete problem $L_h(u_h) = f_h$ has a unique solution for all $h > 0$, and, if $f_h \to f$, then $u_h \to u$.

The first assumption is a formal statement of the intrinsic physical properties (maximum principle, positivity, local bounds, etc.) of the governing operator and the exact solution of (2.1). The second assumption implies that the abstract problem can be discretized and solved in a stable and accurate manner. However, there is no assumption that the discrete solutions $u_h \in \mathscr{U}$, nor is it assumed that $f_h \in \mathscr{V}$, i.e., the discretized problems are not required to preserve the intrinsic physical properties of the original problem.

Optimization-based modeling is a divide and conquer strategy for the preservation of physical properties, which relieves the design of the discrete spaces and operators from that task and assigns it to a reformulation of (2.1) into a suitable *constrained optimization problem.* In so doing our approach allows to separate the accuracy considerations, which are addressed by the discretization spaces and operators, from the enforcement of the physical properties, which are addressed by the optimization formulation.

To explain the main idea, assume for simplicity a conforming approximation setting in which $X_h \subset X$, $Y_h \subset Y$, $\mathscr{U}_h = \mathscr{U} \cap X_h$, and $\mathscr{V}_h = \mathscr{V} \cap Y_h$. The choice of L_h, X_h, and Y_h is governed solely by accuracy considerations and we do not require that the solution of the discrete problem $L_h(u_h) = f_h$ satisfies $u_h \in \mathscr{U}_h$ nor do we assume that $f_h \in \mathscr{V}_h$. Instead, we view the solution of the discrete problem as an *optimization target* $u_h^{\mathsf{T}} := u_h$ and introduce a separate discrete variable $\widehat{u}_h$, which is constrained to have the desired physical properties, i.e., $\widehat{u}_h \in \mathscr{U}_h$ and $L_h(\widehat{u}_h) \in \mathscr{V}_h$. We seek the approximate numerical solution of (2.1) by minimizing the distance, in some normed space W, between $\widehat{u}_h$ and the target, subject to these constraints, i.e., as a solution of the constrained optimization problem

$$\left\{\begin{array}{l} \text{minimize} \quad J(\widehat{u}_h; u_h^{\mathsf{T}}) := \frac{1}{2}\|\widehat{u}_h - u_h^{\mathsf{T}}\|_W^2 \qquad \text{subject to} \\[1em] L_h(u_h^{\mathsf{T}}) = f_h\,; \quad \widehat{u}_h \in \mathscr{U}_h\,; \quad \text{and} \quad L_h(\widehat{u}_h) \in \mathscr{V}_h\,. \end{array}\right. \tag{2.2}$$

In the optimization and control theory this setting is sometimes referred to as the "tracking problem". By solving (2.2) we find the best possible, with respect to the given target u_h^{T}, approximate solution of (2.1), which also possesses the desired physical properties.

The decoupling of the target definition from the task of preserving the physical properties is the hallmark of the optimization approach. Because stability and accuracy are the only requirements for the numerical scheme that defines the target, we can use freely adaptive procedures to resolve regions where the solution experiences rapid change, such as boundary or internal layers. Similarly, if the solution or its derivatives have discontinuities across an interface, the target can be modified to better capture this behavior by using, e.g., discontinuous spaces. Since the preservation of features is not a factor in the definition of the target, the process of adapting the target to the solution can be substantially simplified. In other words, by not requiring the target to be feasible for the optimization problem we significantly simplify the process of selecting a target.

Following a formal statement of the remap problem in Section 3, we specialize (2.2) to a generic feature-preserving optimization-based remap (OBR) algorithm in Section 4 and introduce three specific instances of OBR.

3. Notation, technical background and statement of the remap problem

Let Ω be a polyhedral domain in $\mathbb{R}^d$, $d = 1, 2, 3$, with boundary $\Gamma = \partial\Omega$. We assume that Ω is endowed with two mesh partitions, $C(\Omega)$ and $\widetilde{C}(\Omega)$, respectively. We refer to $C(\Omega)$ to as the "old" or "Lagrangian" mesh, and to $\widetilde{C}(\Omega)$ as the "new" or "rezoned" mesh, respectively[2].

---

[2] This terminology originates with Arbitrary Lagrangian Eulerian methods.

3.1. Mesh entities

This section introduces entities and quantities on the old and the new meshes that are relevant to the formulation of the optimization-based remap. The cells c_i, $i = 1, \ldots, C$, of the old mesh $C(\Omega)$ are d-dimensional closed polytopes[3] defined by a finite number of vertices. The set of all distinct vertices v_i in the mesh is denoted by $V(\Omega)$ and $V = |V(\Omega)|$ is its cardinality. Both convex and non-convex polytopes are admissible as cells. Some methods, such as AMR [35], involve cells with "hanging nodes". We treat these cells as polytopes defined by the totality of their nodes, i.e., a quadrilateral with one hanging node is a pentagon.

The boundary of $c_i \in C(\Omega)$ is a union of $d-1$ dimensional polytopes, called "sides" or "facets" of the cell. We assume that $C(\Omega)$ is a conforming partition in the following sense [36, p.51]: any side s of any cell $c_i \in C(\Omega)$ is either a subset of the boundary Γ, or a side of another cell $c_j \in C(\Omega)$. The set of all distinct sides s_i in the mesh is $S(\Omega)$ and its cardinality is $S = |S(\Omega)|$.

We assume that the new mesh $\widetilde{C}(\Omega)$ has the same connectivity as $C(\Omega)$. Therefore, $\widetilde{C}(\Omega)$ is also a conforming partition. The entities and the quantities on the new mesh will have a tilde accent, whereas the entities and the quantities on $C(\Omega)$ will have no accent. For instance, $\widetilde{c}_i$, $\widetilde{s}_j$ and $\widetilde{v}_k$ denote a cell, side and a vertex on the new mesh. However, because $\widetilde{C}(\Omega)$ and $C(\Omega)$ have the same connectivity, the cardinalities of $\widetilde{C}(\Omega)$, $\widetilde{S}(\Omega)$ and $\widetilde{V}(\Omega)$, match those of $C(\Omega)$, $S(\Omega)$ and $V(\Omega)$, respectively, i.e., $\widetilde{C} = C$, $\widetilde{S} = S$ and $\widetilde{V} = V$. Furthermore, without loss of generality we may assume that the entities in $\widetilde{C}(\Omega)$, $\widetilde{S}(\Omega)$ and $\widetilde{V}(\Omega)$ are ordered in the same way as their Lagrangian prototypes. In other words, the Lagrangian prototypes of $\widetilde{c}_i$, $\widetilde{s}_j$ and $\widetilde{v}_k$ are c_i, s_j and v_k, respectively.

For reasons that will be explained later, we assume that the positions of the boundary vertices in the old and the new meshes are subject to the following two restrictions. First, the vertices defining the polyhedral domain Ω belong to both the old and the new mesh partitions. Second, if the new vertex $\widetilde{v}_i \in \Gamma$, then its Lagrangian prototype $v_i \in \Gamma$, and both $\widetilde{v}_i$ and v_i belong to the same polyhedral part of Γ.

Beside the entities on the old and the new mesh we need some additional entities defined by the movement of the old sides into their new positions. Specifically, the *swept region* r_i, associated with a side $s_i \in S(\Omega)$, is the polytope swept by s_i as it transforms into a side $\widetilde{s}_i \in \widetilde{S}(\Omega)$. Note that the assumption on the positions of the boundary vertices in the old and the new mesh implies that any swept region associated with a boundary side has an empty interior.

The sets of all vertices, sides and cells in an entity $\mathscr{E}$ are $V(\mathscr{E})$, $S(\mathscr{E})$ and $C(\mathscr{E})$, respectively. For instance, $S(c_i)$ are the sides of an old cell c_i, $V(\widetilde{c}_i)$ are the vertices of a new cell $\widetilde{c}_i$, and $S(\Gamma)$ are the sides of the old mesh that are subsets of the domain's boundary. The neighborhood $N(c_i)$ of c_i includes the cell c_i itself and all cells that have a non-empty intersection with c_i. The *side* neighborhood $N_S(c_i)$ comprises of c_i and all cells that share a side with c_i. The *strict* neighborhoods are $N'(c_i) = N(c_i) \setminus c_i$ and $N'_S(c_i) = N(c_i) \setminus c_i$. Note that $|N'_S(c_i)| = |S(c_i)|$, that is the number of cells in the strict side neighborhood of c_i equals the number of sides in c_i. Figure 1 shows an example of $N(c_i)$, $N'(c_i)$, and $N_S(c_i)$ in two dimensions. We assume that $\widetilde{C}(\Omega)$ satisfies the *locality condition*

$$\widetilde{c}_i \subset N(c_i), \quad \text{for all } \; i = 1, \ldots, C\,, \tag{3.1}$$

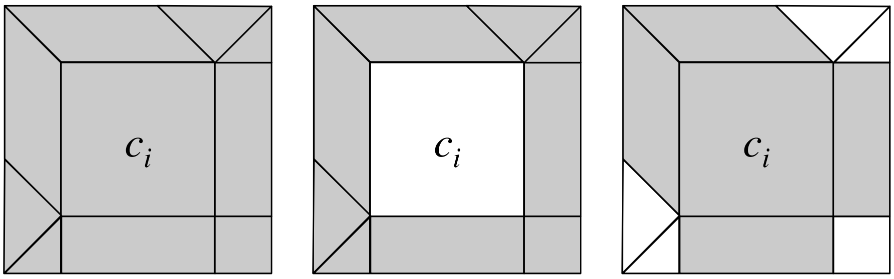
which makes precise the assumption that $C(\Omega)$ and $\widetilde{C}(\Omega)$ are "close".

Suppose that F and $\widetilde{F}$ are some collections of old and new mesh entities, respectively. The notation $i \in F$, $i \in \widetilde{F}$ means that the index i runs over the collection of entity indices in F, resp. $\widetilde{F}$. For instance, $j \in N(c_i)$ indicates that j loops over the cells in the neighborhood of c_i, whereas $k \in S(c_i)$ means that k loops over the sides of c_i.

Lowercase letters with subscripts denote various scalar entities and quantities, whereas the same letter without the subscript denotes the corresponding vector of values. For example, $\widetilde{s}_i$ is a side in the new mesh, whereas $\widetilde{s}$ is the vector of all sides in $\widetilde{C}(\Omega)$. Similarly, if m_i is the mass on the old cell c_i, then m is a vector of all cell masses.

---

[3] For instance, in two dimensions c_i can be triangles, quadrilaterals or polygons (2-polytopes), whereas in three dimensions the admissible cells are tetrahedrons, hexahedrons or polyhedrons (3-polytopes).



$N(c_i)$ $N'(c_i)$ $N_S(c_i)$

Figure 1: Lagrangian cell c_i, the neighborhood $N(c_i)$, the strict neighborhood $N'(c_i)$, and the side neighborhood $N_S(c_i)$.

3.2. Mesh spaces

It is convenient to think of the quantities involved in the remap problem as elements of approximation spaces defined on the old and the new meshes, and associated with specific mesh entities. The space C^h is the piecewise constant space defined with respect to the cells in the old mesh $C(\Omega)$ and C^h_0 is its "zero-mean" subspace:

$$C^h_0 = \Big\{ u \in C^h \mid \sum_{i=1}^{C} u_i = 0 \Big\} .$$

The analogues of these spaces on the new mesh are $\widetilde{C}^h$ and $\widetilde{C}^h_0$, respectively. The elements of these spaces are vectors in $\mathbb{R}^C$, indexed by a cell number. The value of each entry represents a *cell-centered* quantity such as the mean cell value of a given function.

The space S^h is a piecewise constant space defined with respect to the sides $S(\Omega)$ on the old mesh, and S^h_0 is the subspace of all elements in S^h that satisfy a homogeneous "boundary condition"

$$S^h_0 = \Big\{ u \in S^h \mid u_i = 0 \quad \forall i \in S(\Gamma) \Big\} ,$$

i.e., their values associated with boundary sides are zero. The new mesh analogues of these spaces are $\widetilde{S}^h$ and $\widetilde{S}^h_0$, respectively. The elements of S^h and $\widetilde{S}^h$ are vectors in $\mathbb{R}^S$, indexed by a side number. Their entries are *side-centered* quantities associated with the sides, such as fluxes through the sides or integrals of functions over the swept regions.

We endow C^h and S^h with the standard Euclidean inner product and norm, that is,

$$\|u\|_C = \Big(\sum_{i=1}^{C} u_i^2 \Big)^{1/2} \quad \forall u \in C^h \quad \text{and} \quad \|v\|_S = \Big(\sum_{i=1}^{S} v_i^2 \Big)^{1/2} \quad \forall v \in S^h , \tag{3.2}$$

respectively.

3.3. Mesh entity orientations and measures

Computation of quantities relevant to the remap requires orientations of cells, sides and swept regions, as well as a notion of a signed measure of a domain. For clarity, we explain the orientation choices in two-dimensions.

The direction of the unit normal to the boundary of a cell specifies the orientation of that cell. Let $\boldsymbol{n}^c_i$ and $\widetilde{\boldsymbol{n}}^c_i$ denote the outer unit normals to $c_i \in C(\Omega)$ and $\widetilde{c}_i \in \widetilde{C}(\Omega)$, respectively. We use these unit normals to orient the cells in the old and the new mesh; see the left pane in Figure 2. The direction of the unit normal to a side orients the side. If a side c_k belongs to $\partial\Omega$, we use the outer normal to $\partial\Omega$. On the other hand, if a side c_k is in the interior, i.e., $s_k = c_i \cap c_j$ for some cells c_i and c_j, we choose its unit normal $\boldsymbol{n}^s_k$ according to the following rule:

$$\boldsymbol{n}^s_k = \begin{cases} \boldsymbol{n}^c_i & \text{if } i < j \\ \boldsymbol{n}^c_j & \text{if } i > j \, . \end{cases}$$

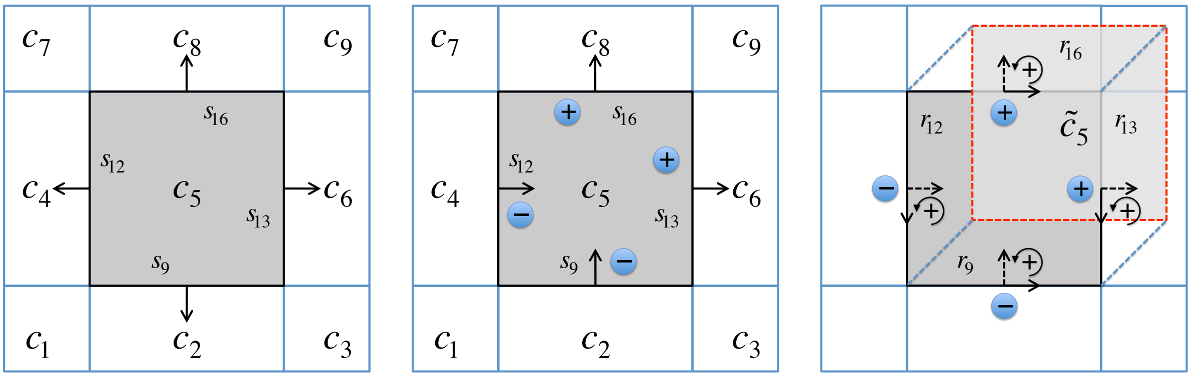


Figure 2: Example orientations in two dimensions. The cells are oriented by the outer unit normal to their boundary (left pane). Sides are oriented by the outer unit normal of the cell with the smaller index (center pane). The blue ovals are the signs of the non-zero entries in the row of the side-to-cell incidence matrix (3.4) corresponding to c_5. Swept regions are oriented by the unit tangent for which $\{\boldsymbol{t}^s_k, \boldsymbol{n}^s_k\}$ is positively oriented (right pane). The combination of the signs in the blue ovals and the orientations of the swept regions illustrates formula (3.10).

The center pane in Figure 2 illustrates this rule. For example, the side $s_9 = c_2 \cap c_5$ is oriented by the outer normal to c_2, whereas $s_{13} = c_5 \cap c_6$ is oriented by the outer normal to c_5. This rule applies to both the old and the new sides.

In two dimensions we need to compute contour integrals over the swept regions. The orientations of the swept regions require the selection of unit tangents on their boundaries. Suppose that r_k is the swept region corresponding to the side s_k and $\boldsymbol{n}^s_k$ is the unit normal that specifies the side's orientation. We orient r_k using the unit tangent $\boldsymbol{t}^s_k$ for which the pair $\{\boldsymbol{t}^s_k, \boldsymbol{n}^s_k\}$ is positively oriented, that is, $\det([\boldsymbol{t}^s_k, \boldsymbol{n}^s_k]) = 1$. The right pane in Figure 2 shows the orientations of the swept regions formed by the movement of the sides of cell c_5 into their new positions in $\widetilde{c}_5$.

The unsigned measure of a set $\mathscr{D}$ is

$$\mu(\mathscr{D}) = \int_{\mathscr{D}} dV\,.$$

The reduction of the volume integral to an integral over an oriented surface defines the signed measure $\mu^\star(\mathscr{D})$ of the set. For instance, in two-dimensions, the Green's Theorem implies that

$$\int_{\mathscr{D}} dxdy = \int_{\partial\mathscr{D}} xdy = -\int_{\partial\mathscr{D}} ydx\,.$$

The specification of an orientation on $\partial\mathscr{D}$ yields the signed measure $\mu^\star(\mathscr{D})$. We recall the barycenter formula for the old and the new cells

$$\boldsymbol{b}_j = \frac{\displaystyle\int_{c_j} \boldsymbol{x}\, dV}{\mu(c_j)}\,, \quad \text{and} \quad \widetilde{\boldsymbol{b}}_j = \frac{\displaystyle\int_{\widetilde{c}_j} \boldsymbol{x}\, dV}{\widetilde{\mu}(c_j)}\,, \tag{3.3}$$

where $\boldsymbol{x}$ is the position vector in $\mathbb{R}^d$.

3.4. The side-to-cell incidence matrix

The side-to-cell incidence matrix $\mathbf{D}$ for $C(\Omega)$ is a $C \times S$ matrix with entries $D_{ij} \in \{0, -1, 1\}$ such that

$$D_{ij} = \begin{cases} 0 & \text{if } s_j \notin S(c_i) \\ 1 & \text{if } s_j \in S(c_i) \text{ and } \boldsymbol{n}^s_j = \boldsymbol{n}^c_i \\ -1 & \text{if } s_j \in S(c_i) \text{ and } \boldsymbol{n}^s_j = -\boldsymbol{n}^c_i \end{cases}. \tag{3.4}$$

In other words, the value of D_{ij} indicates whether or not the side s_j belongs to the cell c_i, and if it does – the direction of its normal $\boldsymbol{n}^s_j$ relative to the outer normal on c_i. The blue ovals in Figure 2 illustrate this rule. The entries of $\mathbf{D}$ depend only on the mesh topology and not on the shape of the cells themselves. As a result, $\widetilde{\mathbf{D}} = \mathbf{D}$. In algebraic topology the matrix $\mathbf{D}$ represents the coboundary operator acting on 2-cochains. This operator gives rise to compatible discretizations of the divergence [37], [38] in numerical methods for PDEs. The connection between $\mathbf{D}$ and the divergence is responsible for many interesting properties of this matrix. A summary of the properties relevant to remap follows. Proofs of some of these properties rely on fundamental results from algebraic topology and exterior calculus and are beyond the scope of this paper. We refer to [5, 38, 37, 39] and the references therein for further details.

Full row rank property. The matrix $\mathbf{D}$ has full row rank. This property follows from the fact that the non zero elements in the ith row of $\mathbf{D}$ correspond to the sides that belong to the boundary of the ith cell in the mesh. Suppose that two rows are linearly dependent. This would imply that the boundaries of two distinct cells comprise of exactly the same sides, which is impossible, as two cells share at most a single side.

Surjective property. Considered as a mapping $S^h_0 \mapsto C^h_0$, the matrix $\mathbf{D}$ is a surjection:[4]

$$\forall m \in C^h_0 \quad \exists u_m \in S^h_0 \quad \text{such that} \quad m = \mathbf{D} u_m \,. \tag{3.5}$$

It is easy to see that the range of $\mathbf{D}$ is a subset of C^h_0, i.e.,

$$\sum_{i=1}^{C} (\mathbf{D}u)_i = 0 \quad \forall u \in S^h_0 \,. \tag{3.6}$$

Indeed, all entries of $u \in S^h_0$ associated with the boundary sides equal zero and contribute nothing to the sum. On the other hand, if an entry u_j of u corresponds to an interior side s_j, this side is shared by exactly two cells, and the corresponding column of $\mathbf{D}$ has exactly two nonzero entries. Because the orientation of s_j matches the orientation of one of the cells and is opposite the orientation of the other cells, the column entries are $+1$ and -1, respectively. Therefore, every element u_j associated with interior sides enters the sum in (3.6) exactly twice with opposite signs. The proof of the converse, i.e., that any "zero-mean" cell-centered quantity $m \in C^h_0$ is an image, under the action of $\mathbf{D}$, of a side-centered quantity $u \in S^h_0$ with homogeneous "boundary condition", follows from the full row rank property of $\mathbf{D}$.

Null space. The matrix $\mathbf{D}$ has a nontrivial null space, $\ker \mathbf{D}$, in $\mathbb{R}^d$, $d = 1, 2, 3$. For example, in three dimensions the columns of the *edge-to-side* incidence matrix $\mathbf{C}$ span $\ker \mathbf{D}$, i.e., $\mathbf{DC} = 0$.[5]

Orthogonal decomposition of S^h_0. The space S^h_0 admits an orthogonal decomposition that mimics the properties of the classical Hodge decomposition of vector fields. Specifically, we have that

$$S^h_0 = \ker \mathbf{D} \oplus \ker \mathbf{D}^\perp \,, \tag{3.7}$$

see [38], i.e., every element $u \in S^h_0$ has the form $u = u^0 + u^\perp$, where $u^0 \in \ker \mathbf{D}$ and $u^\perp \in \ker \mathbf{D}^\perp$, the orthogonal complement. In addition, a discrete Poincare-Friedrichs inequality holds on the null-space complement: under some assumptions on the shape of the mesh cells, there is a positive constant K, which is independent of the mesh size, such that

$$\|u^\perp\|_S \le K \|\mathbf{D}u\|_C \quad \forall u^\perp \in \ker \mathbf{D}^\perp \,. \tag{3.8}$$

The same properties hold on $\widetilde{S}^h_0$.

---

[4] This property is a discrete version of the surjective property of the divergence operator as a mapping $H_0(div, \Omega) \mapsto L^2_0(\Omega)$.

[5] In algebraic topology $\mathbf{C}$ is the coboundary operator acting on 1-cochains. This operator gives rise to the discretization of the curl operator in compatible discretization methods for PDEs [38]. The property $\mathbf{DC} = 0$ mimics the vector calculus identity $\operatorname{div} \operatorname{curl} \mathbf{u} = 0$.

Boundary and measure representations. For any cell c_i there holds[6]

$$\partial c_i = \sum_{j \in S(c_i)} D_{ij} s_j \,. \tag{3.9}$$

Property (3.9) holds on the new mesh as well. Using the vector notation formula (3.9) assumes the compact form $\partial c = \mathbf{D}s$. Using the relationship (3.9) one can show that with the orientation choices in Section 3.3

$$\mu(\widetilde{c}_i) = \mu(c_i) + \sum_{j \in S(c_i)} D_{ij} \mu^\star(r_j) \,. \tag{3.10}$$

We can write (3.10) in a compact matrix form as

$$\mu(\widetilde{c}) = \mu(c) + \mathbf{D}\mu^\star(r) \,.$$

Note that the assumptions on the positions of the boundary vertices imply that $\mu^\star(r) \in S^h_0$, which, in conjunction with the surjective property of $\mathbf{D}$ implies that $\mathbf{D}\mu^\star(r) \in C^h_0$. Therefore,

$$\sum_{i=1}^{C} \mu(\widetilde{c}_i) = \sum_{i=1}^{C} \Big(\mu(c_i) + (\mathbf{D}\mu^\star(r))_i \Big) = \sum_{i=1}^{C} \mu(c_i) \,,$$

which is simply a statement of the fact that the measure of the computational domain does not depend on the mesh. This property motivates the restrictions on the positions of the boundary nodes, stated earlier.

3.5. Statement of the mass-density remap problem

We state the mass-density remap problem following [25, 30]. The primitive variable (density) is a scalar function $\rho(\boldsymbol{x}) > 0$ on Ω and the conserved variable is the total mass

$$M = \int_\Omega \rho(\boldsymbol{x}) dV \,.$$

The cell masses and the mass densities on the old mesh are

$$m_i = \int_{c_i} \rho(\boldsymbol{x}) dV \quad \text{and} \quad \rho_i = \frac{\displaystyle\int_{c_i} \rho(\boldsymbol{x}) dV}{\mu(c_i)} \,, \quad i = 1, 2, \ldots, C \,, \tag{3.11}$$

respectively. Therefore,

$$m_i = \rho_i \mu(c_i) \,, \qquad \rho_i = \frac{m_i}{\mu(c_i)} \qquad \text{and} \quad M = \sum_{i=1}^{C} \rho_i \mu(c_i) \,.$$

For every old cell c_i we define

$$\rho_i^{\min} = \begin{cases} \displaystyle\min_{j \in N(c_i)} \{\rho_j\} & \text{if } c_i \cap \partial\Omega = \emptyset \\ \min\Big\{ \displaystyle\min_{j \in N(c_i)} \{\rho_j\}, \min_{\boldsymbol{x} \in N(c_i) \cap \partial\Omega} \rho(\boldsymbol{x}) \Big\} & \text{if } c_i \cap \partial\Omega \neq \emptyset \end{cases} \,, \tag{3.12}$$

$$\rho_i^{\max} = \begin{cases} \displaystyle\max_{j \in N(c_i)} \{\rho_j\} & \text{if } c_i \cap \partial\Omega = \emptyset \\ \max\Big\{ \displaystyle\max_{j \in N(c_i)} \{\rho_j\}, \max_{\boldsymbol{x} \in N(c_i) \cap \partial\Omega} \rho(\boldsymbol{x}) \Big\} & \text{if } c_i \cap \partial\Omega \neq \emptyset \,. \end{cases} \tag{3.13}$$

---

[6] This property reflects the fact that the coboundary $\mathbf{D}$ is dual to the boundary operator.

It is straightforward to check that the mean density on the old cells satisfies the local bounds

$$\rho_i^{\min} \leq \rho_i \leq \rho_i^{\max}\,, \quad i = 1, \ldots, C\,. \tag{3.14}$$

The inequalities (3.14) together with (3.11) imply that the old cell masses satisfy similar local bounds:

$$m_i^{\min} := \rho_i^{\min}\mu(c_i) \leq m_i \leq \rho_i^{\max}\mu(c_i) =: m_i^{\max}\,, \quad i = 1, \ldots, C\,. \tag{3.15}$$

In the mass-density remap problem we are given the mean density values ρ_i on the *old* mesh cells c_i and the boundary values of $\rho(\boldsymbol{x})$ on $\partial\Omega$. The objective is to find accurate approximations $\widetilde{m}_i$ for the masses of the *new* cells $\widetilde{c}_i$,

$$\widetilde{m}_i \approx \widetilde{m}_i^{\mathrm{ex}} = \int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV\,, \quad i = 1, \ldots, C\,. \tag{3.16}$$

We require that the approximations of the new masses satisfy the following conditions:

C1. The total mass is conserved:

$$\sum_{i=1}^{C} \widetilde{m}_i = \sum_{i=1}^{C} m_i = M\,. \tag{3.17}$$

C2. If $\rho(\boldsymbol{x})$ is a global linear function on Ω, then the remapped masses are exact:

$$\widetilde{m}_i = \widetilde{m}_i^{\mathrm{ex}} = \int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV\,, \quad i = 1, \ldots, C\,. \tag{3.18}$$

C3. The approximation of the mean density on the new cells

$$\widetilde{\rho}_i = \frac{\widetilde{m}_i}{\widetilde{\mu}_i} \tag{3.19}$$

satisfies the local bounds

$$\rho_i^{\min} \leq \widetilde{\rho}_i \leq \rho_i^{\max}\,; \quad i = 1, \ldots, C\,, \tag{3.20}$$

where (3.12)–(3.13) define $\rho_i^{\min}$ and $\rho_i^{\max}$. Equivalently, the following local mass bounds hold:

$$\widetilde{m}_i^{\min} := \rho_i^{\min}\widetilde{\mu}_i \leq \widetilde{m}_i \leq \rho_i^{\max}\widetilde{\mu}_i =: \widetilde{m}_i^{\max}\,. \tag{3.21}$$

□

4. Optimization-based formulations of remap

This section applies the optimization-based approach in Section 2 to the solution of the mass-density remap in Section 3.5. We first specialize the abstract optimization problem (2.2) to a generic optimization-based formulation of remap (OBR) and then specialize this formulation to specific instances of OBR.

The key step in remap is the transformation of the given data on the old mesh into approximate cell masses $\widetilde{m} \in \widetilde{C}^h$ on the new mesh such that (3.17) and (3.21) hold. To formalize this step, assume that $\rho(\boldsymbol{x})$ belongs in some function space R and let $L : R \mapsto \widetilde{C}^h$ be an *exact mass update* operator

$$\widetilde{m}^{\mathrm{ex}} = L(\rho(\boldsymbol{x}))\,. \tag{4.1}$$

In general, we cannot compute the action of L because the exact density function $\rho(\boldsymbol{x})$ is not known and we are only given the old cell masses $m \in C^h$ and the old mean cell densities $\rho \in C^h$. Therefore, we replace L by an approximation $L_h : C^h \times C^h \mapsto \widetilde{C}^h$, which operates on the given data. We seek the approximate operator in the form $L_h(m, u(\rho))$ where $u(\rho)$ is some quantity that depends on the mean density values and belongs in a space $\mathscr{E}^h$, associated with some mesh entity type $\mathscr{E}$. A subset $\mathscr{U}^h \subseteq \mathscr{E}^h$ defines the "physical" properties of $u(\rho)$. On the other hand, the conservation of the total mass (3.17) and the local bounds (3.21) define the constraints $\mathscr{V}_h$ on the range of L_h.

In the optimization-based framework the only consideration in the definition of L_h is the accuracy of the new cell masses. Therefore, we disregard the range constraints $\mathscr{V}_h$ and require that $L_h(m, u(\rho))$ is exact for a given class $\mathscr{C}$ of scalar density functions:

$$\widetilde{m}^{\text{ex}} = L_h(m, u(\rho)) \qquad \forall \rho(\boldsymbol{x}) \in \mathscr{C}\,. \tag{4.2}$$

For example, if $\mathscr{C}$ is the class of globally linear scalar functions, condition (4.2) requires $L_h(m, u(\rho))$ to compute the exact new cell masses whenever $\rho(\boldsymbol{x})$ is a linear function. However, the output of L_h is not required to conserve the total mass, nor is it required to satisfy the local bounds (3.21). To define the optimization targets we treat (4.2) as an equation for $u(\rho)$, i.e., the target u^{T} is the element of $\mathscr{E}_h$ such that

$$L_h(m, u^{\mathsf{T}}) = \widetilde{m}^{\text{ex}} \qquad \forall \rho(\boldsymbol{x}) \in \mathscr{C}\,.$$

Taking all of the above into consideration and setting[7] $W = \ell_2$, we specialize (2.2) to the generic OBR formulation

$$\begin{cases} \underset{\widehat{u} \in \mathscr{U}^h}{\text{minimize}} \quad \dfrac{1}{2}\|\widehat{u} - u^{\mathsf{T}}\|^2_{\ell_2} \qquad \text{subject to} \\ L_h(m, u^{\mathsf{T}}) = \widetilde{m}^{\text{ex}} \quad \forall \rho(\boldsymbol{x}) \in \mathscr{C}\,; \\ \displaystyle\sum_{i=1}^{C} m_i = \sum_{i=1}^{C} (L_h(m, \widehat{u}))_i \quad \text{and} \quad \widetilde{m}^{\min} \leq L_h(m, \widehat{u}) \leq \widetilde{m}^{\max}\,. \end{cases} \tag{4.3}$$

The generic OBR problem (4.3) is a quadratic program (QP) in which the objective is to match a given target quantity, subject to physically motivated constraints on the range of the approximate mass update operator. The selection of a specific approximate mass update operator and an optimization objective specializes (4.3) further. In what follows we consider three different specializations of (4.3).

4.1. The flux-variable flux-target (FVFT) optimization formulation of remap

This section specializes the generic OBR problem (4.3) to a *flux-variable flux-target* (FVFT) formulation. The starting point is the representation of the new cell masses in the *mass flux* form

$$\int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV = \int_{c_i} \rho(\boldsymbol{x}) dV + \sum_{j \in S(c_i)} D_{ij} \int_{r_j} \rho(\boldsymbol{x}) dV\,, \tag{4.4}$$

which follows from formula (3.10). The right hand side in (4.4) defines an exact mass update operator $L : R \mapsto \widetilde{C}^h$. We seek the approximate mass update operator in the form

$$\widetilde{m} = L_h(m, u(\rho)) := m + \mathbf{D}u(\rho)\,, \tag{4.5}$$

where the quantity $u(\rho) \in \mathscr{E}^h$ approximates the exact *mass fluxes*,

$$u_j^{\text{ex}} = \int_{r_j} \rho(\boldsymbol{x}) dV \quad j \in S(\Omega)\,, \tag{4.6}$$

on the swept regions. Because a swept region is associated with a side, the entities $\mathscr{E}$ are the mesh sides $S(\Omega)$ and $\mathscr{E}^h$ is the space S^h. Owing to the assumptions on the positions of the boundary vertices, we recall that $\mu(r_j) = 0$ whenever r_j corresponds to a boundary side. As a result, $u_j^{\text{ex}} = 0$ for all $j \in S(\Gamma)$, i.e., $u^{\text{ex}} \in S_0^h$. This allows us to identify $\mathscr{U}^h$ with the imposition of a homogeneous “boundary condition” on the approximate fluxes $u(\rho)$.

---

[7] We choose this setting for simplicity because it yields differentiable objective functions. Choices such as $W = \ell_p$ are also possible, but the resulting optimization problems may be more difficult to solve.

Furthermore, since $\mathbf{D}$ maps S_0^h into the zero mean space C_0^h, the imposition of the homogeneous boundary condition on $u(\rho)$ guarantees that the FVFT mass update operator conserves the total mass,

$$\sum_{i=1}^{C} \widetilde{m}_i = \sum_{i=1}^{C} (L_h(m, u(\rho)))_i = \sum_{i=1}^{C} m_i + (\mathbf{D}u(\rho))_i = \sum_{i=1}^{C} m_i \quad \forall u(\rho) \in S_0^h ,$$

that is, the restriction of $u(\rho)$ to S_0^h, in conjunction with (4.5), obviates the equality constraint.

In this paper we restrict attention to linearity preserving remap. Therefore, to define the target fluxes we take $\mathscr{C}$ to be the class of globally linear scalar functions and seek $u^{\mathsf{T}} \in S^h$ such that

$$L_h(m, u^{\mathsf{T}}) \equiv m + \mathbf{D}u^{\mathsf{T}} = \widetilde{m}^{\mathrm{ex}}_{\mathscr{C}} \qquad \forall \rho \in \mathscr{C} . \tag{4.7}$$

It is straightforward to check that (4.7) holds if

$$u_j^{\mathsf{T}} := \int_{r_j} \rho^h(\boldsymbol{x}) dV , \tag{4.8}$$

for some density reconstruction $\rho^h(\boldsymbol{x})$ that is exact for linear functions. Note that while the target is not subject to any physical constraints it nonetheless satisfies $u^{\mathsf{T}} \in S_0^h$ and $\widetilde{m}^{\mathsf{T}} := L_h(m, u^{\mathsf{T}})$ conserves the total mass. However, the masses $\widetilde{m}^{\mathsf{T}}$ are not guaranteed to satisfy the physically motivated local bounds.

Taking all of the above into consideration, the generic OBR problem (4.3) yields the following FVFT formulation:

$$\begin{cases} \underset{\widehat{u} \in S_0^h}{\text{minimize}} \quad \dfrac{1}{2} \| \widehat{u} - u^{\mathsf{T}} \|_S^2 \qquad \text{subject to} \\ u_j^{\mathsf{T}} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x}) dV, \quad j = 1, \dots, S ; \\ \widetilde{m}^{\min} \leq m + \mathbf{D}\widehat{u} \leq \widetilde{m}^{\max} . \end{cases} \tag{4.9}$$

In practice, one enforces the boundary condition constraint $\widehat{u} \in S_0^h$ directly on the candidate minimizers, which reduces the FVFT problem (4.9) a global inequality-constrained QP.

4.2. The mass-variable mass-target (MVMT) optimization formulation of remap

The conservation of the total mass by the FVFT mass update operator (4.5) when $u(\rho)$ satisfies a homogeneous boundary condition represents an attractive computational property of FVFT. On the other hand, this operator involves the action of $\mathbf{D}$ on elements of S_0^h, which results in a globally coupled system of linear inequality constraints. This coupling complicates the solution of (4.9) and creates a potential performance bottleneck.

In this section we specialize the generic OBR problem (4.3) into an alternative *mass-variable mass-target* (MVMT) formulation. Instead of (4.4), the starting point is the representation of the new cell masses in the *incremental mass* form

$$\int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV = \int_{c_i} \rho(\boldsymbol{x}) dV + \left(\int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV - \int_{c_i} \rho(\boldsymbol{x}) dV \right) . \tag{4.10}$$

Again, the right-hand side in (4.10) defines an exact mass update operator $L : R \mapsto \widetilde{C}^h$. We approximate this operator by

$$\widetilde{m} = L_h(m, u(\rho)) := m + u(\rho) , \tag{4.11}$$

where the quantity $u(\rho) \in \mathscr{E}^h$ approximates the exact *mass increments* between the new cells and the old cells,

$$u_i^{\mathrm{ex}} = \int_{\widetilde{c}_i} \rho(\boldsymbol{x}) dV - \int_{c_i} \rho(\boldsymbol{x}) dV \quad i \in C(\Omega) . \tag{4.12}$$

Because the exact mass increments are associated with the cell indices, the entities $\mathscr{E}$ are the mesh cells $C(\Omega)$, and $\mathscr{E}^h$ is the space C^h. Owing to the assumption on the positions of the boundary vertices

$$\sum_{i=1}^{C} u_i^{\text{ex}} = \sum_{i=1}^{C} \int_{\widetilde{c}_i} \rho(\boldsymbol{x})dV - \sum_{i=1}^{C} \int_{c_i} \rho(\boldsymbol{x})dV = \int_{\Omega} \rho(\boldsymbol{x})dV - \int_{\Omega} \rho(\boldsymbol{x})dV = 0\,. \tag{4.13}$$

It follows that $u^{\text{ex}} \in C_0^h$, which allows us to identify $\mathscr{U}^h$ with the imposition of a zero mean constraint on the cell values.

Furthermore, the imposition of this constraint on the approximate mass increments $u(\rho)$ guarantees that the MVMT mass update operator conserves the total mass:

$$\sum_{i=1}^{C} \widetilde{m}_i = \sum_{i=1}^{C} (L(m, u(\rho)))_i = \sum_{i=1}^{C} m_i + u(\rho)_i = \sum_{i=1}^{C} m_i \quad \forall u(\rho) \in C_0^h\,.$$

This property is reminiscent of the situation encountered in Section 4.1, where the imposition of the homogeneous "boundary condition" on the mass fluxes provided for the conservation of the total mass, but for one important distinction. Restriction of an element $u \in S^h$ to S_0^h is trivial because the "boundary" values of u can be set to zero independently of each other. On the other hand, forcing an element $u \in C^h$ to have a zero mean is equivalent to solving a single linear equation which couples *all* cell values and has infinitely many solutions. Therefore, to guarantee mass conservation we must enforce $u \in C_0^h$ by explicitly including this linear constraint into the optimization formulation.

As in Section 4.1, we define the optimization target u^{T} by taking $\mathscr{C}$ to be the class of globally linear functions and requiring that

$$L_h(m, u^{\mathsf{T}}) \equiv m + u^{\mathsf{T}} = \widetilde{m}^{\text{ex}}_{\mathscr{C}} \qquad \forall \rho \in \mathscr{C}\,. \tag{4.14}$$

It is easy to see that this equation holds with

$$u_i^{\mathsf{T}} = \int_{\widetilde{c}_i} \rho^h(\boldsymbol{x})dV - \int_{c_i} \rho^h(\boldsymbol{x})dV\,, \tag{4.15}$$

provided the density reconstruction $\rho^h(\boldsymbol{x})$ is exact for linear functions. Using the same arguments as in (4.13) it easily follows that the target has zero mean, i.e., $u^{\mathsf{T}} \in C_0^h$. Thus, $\widetilde{m}^{\mathsf{T}} := L_h(m, u^{\mathsf{T}})$ conserves the total mass but $\widetilde{m}^{\mathsf{T}}$ are not guaranteed to satisfy the physically motivated local bounds.

To summarize, the application of the approximate mass update operator (4.11) specializes the generic OBR formulation (4.9) to the MVMT formulation:

$$\begin{cases} \underset{\widehat{u} \in C^h}{\text{minimize}} \quad \dfrac{1}{2}\|\widehat{u} - u^{\mathsf{T}}\|_C^2 \qquad \text{subject to} \\ u_i^{\mathsf{T}} := \displaystyle\int_{\widetilde{c}_i} \rho^h(\boldsymbol{x})dV - \int_{c_i} \rho^h(\boldsymbol{x})dV, \quad i = 1, \ldots, C\,; \\ \displaystyle\sum_{i=1}^{C} \widehat{u}_i = 0 \quad \text{and} \quad \widetilde{m}^{\min} \le m + \widehat{u} \le \widetilde{m}^{\max}\,. \end{cases} \tag{4.16}$$

While MVMT requires the equality constraint for the conservation of mass, the inequality constraints in (4.16) are completely decoupled. Thus, switching from the FVFT mass update to the MVMT mass update changes the type of the OBR formulation from a globally inequality-constrained QP to a singly linearly constrained QP with simple bounds. The latter structure lends itself to fast and efficient parallel solution, which makes it particularly attractive for practical algorithm development [2].

4.3. The flux-variable mass-target (FVMT) optimization formulation of remap

Recall the approximate MVMT mass update operator

$$L_h(m, u(\rho)) = m + u(\rho)$$

where $u(\rho) \in C_0^h$. Because $\mathbf{D} : S_0^h \mapsto C_0^h$ is a surjection, for every $u \in C_0^h$ there exists $v_u \in S_0^h$ such that $u = \mathbf{D} v_u$. Furthermore, it is easy to see that the identity

$$\int_{\widetilde{c}_i} \rho^h(\boldsymbol{x}) dV - \int_{c_i} \rho^h(\boldsymbol{x}) dV = \sum_{j \in S(c_i)} D_{ij} \int_{r_j} \rho^h(\boldsymbol{x}) dV$$

holds for any given density reconstruction $\rho^h(\boldsymbol{x})$. The right hand side in this identity involves the FVFT target, which we label $u_F^{\mathsf{T}} \in S_0^h$ below, whereas the left hand side defines the zero-mean MVMT target, denoted by $u_M^{\mathsf{T}} \in C_0^h$. These targets of the two OBR formulations are equivalent in the sense that

$$u_M^{\mathsf{T}} = \mathbf{D} u_F^{\mathsf{T}} . \tag{4.17}$$

Therefore, we can "change" the variables and the targets in the MVMT formulation from mass increments to mass fluxes. In so doing we effectively combine the approximate FVFT mass update operator with the MVMT objective functional. Since the FVFT mass update conserves the mass, the equality constraint in (4.16) is redundant and can be removed from the formulation. The result is the following "hybrid", FVMT problem:

$$\begin{cases} \underset{\widehat{u} \in S_0^h}{\text{minimize}} \quad \frac{1}{2} \| \mathbf{D}\widehat{u} - \mathbf{D} u^{\mathsf{T}} \|_C^2 \qquad \text{subject to} \\ u_j^{\mathsf{T}} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x}) dV, \quad j = 1, \dots, S\,; \\ \widetilde{m}^{\min} \le m + \mathbf{D}\widehat{u} \le \widetilde{m}^{\max} . \end{cases} \tag{4.18}$$

While FVMT combines features of both FVFT and MVMT it is equivalent to neither of the two formulations. For instance, the objectives of both (4.9) and (4.16) are strictly convex, whereas the objective of (4.18) is not. This follows from the fact that $\mathbf{D}$ has a non-trivial kernel; see Section 3.3. Besides the non-strict convexity of the objective, FVMT also inherits the globally coupled inequality constraints of FVFT.

5. Analytical properties of OBR formulations

This section presents recent and new theoretical results concerning the mathematical properties of the OBR formulations. We begin with a theorem that establishes the well-posedness of the FVFT and the MVMT and states sufficient conditions for the preservation of linear densities. The theorem combines several results from [1], [24] and [2] and we refer to these papers for detailed proofs.

Theorem 5.1. *Assume that the locality condition (3.1) holds for $C(\Omega)$ and $\widetilde{C}(\Omega)$ and that $\rho \in C^h$ is the mean cell density on the Lagrangian mesh. Then, the quadratic programs (4.9) and (4.16) have unique optimal solutions.*

Furthermore, assume that the density reconstruction $\rho^h(\boldsymbol{x})$ is exact for linear functions and (4.8) and (4.15) define the targets u_F^T and u_M^T, respectively. Let B_i denote the set of barycenters of the old cells in $N(c_i)$, $B_i = \{\boldsymbol{b}_j \,|\, j \in N(c_i)\}$, and let $\widetilde{\boldsymbol{b}}_i$ be the barycenter of the new cell $\widetilde{c}_i$. The conditions

$$\widetilde{\boldsymbol{b}}_i \in \mathscr{H}(B_i) \qquad \text{if } c_i \cap \partial\Omega = \emptyset, \tag{5.1}$$

$$\widetilde{\boldsymbol{b}}_i \in \mathscr{H}\left(B_i \cup (N(c_i) \cap \partial\Omega)\right) \qquad \text{if } c_i \cap \partial\Omega \neq \emptyset, \tag{5.2}$$

where $\mathscr{H}(\cdot)$ denotes the convex hull, are sufficient for FVFT and the MVMT to yield the exact masses on the new cells, whenever $\rho(\boldsymbol{x}) = c_0 + \mathbf{c} \cdot \boldsymbol{x}$. □

5.1. The structure of the FVFT optimal solution

From (3.7) it follows that the FVFT target $u^{\mathsf{T}} \in S_0^h$ has the orthogonal decomposition

$$u^{\mathsf{T}} = u_0^{\mathsf{T}} + u_\perp^{\mathsf{T}} , \quad u_0^{\mathsf{T}} \in \ker \mathbf{D} \quad \text{and} \quad u_\perp^{\mathsf{T}} \in \ker \mathbf{D}^\perp .$$

Using this decomposition the FVFT objective splits into two independent terms:

$$\|\widehat{u} - u^{\mathsf{T}}\|_S^2 = \|\widehat{u}_0 - u_0^{\mathsf{T}}\|_S^2 + \|\widehat{u}_\perp - u_\perp^{\mathsf{T}}\|_S^2 \,,$$

where $\widehat{u} = \widehat{u}_0 + \widehat{u}_\perp$ is the orthogonal decomposition of $u \in S_0^h$. Because $\mathbf{D}\widehat{u} = \mathbf{D}\widehat{u}_\perp$, the inequality constraints in (4.9) do not depend on the null-space component $\widehat{u}_0$. It follows that the FVFT splits equivalently into two independent optimization problems over $\ker \mathbf{D}^\perp$ and $\ker \mathbf{D}$, respectively,

$$(a)\begin{cases} \underset{\widehat{u}_\perp \in \ker \mathbf{D}^\perp}{\text{minimize}} \quad \dfrac{1}{2}\|\widehat{u}_\perp - u_\perp^{\mathsf{T}}\|_S^2 \quad \text{subject to} \\ u_j^{\mathsf{T}} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x})dV, \quad j = 1,\dots,S\,; \\ \widetilde{m}^{\min} \le m + \mathbf{D}\widehat{u}_\perp \le \widetilde{m}^{\max} \end{cases} \quad \text{and} \quad (b)\begin{cases} \underset{\widehat{u}_0 \in \ker \mathbf{D}}{\text{minimize}} \quad \dfrac{1}{2}\|\widehat{u}_0 - u_0^{\mathsf{T}}\|_S^2 \quad \text{subject to} \\ u_j^{\mathsf{T}} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x})dV, \quad j = 1,\dots,S \end{cases} \tag{5.3}$$

for the orthogonal components of the FVFT optimal solution. Problem (b) is a global QP with a strictly convex objective. The unique minimizer of this problem is the null-space component u_0^{T} of the FVFT target u^{T}. The feasible set of Problem (a) has the form

$$\Sigma_\perp = \{u_\perp \in \ker \mathbf{D}^\perp \,|\, \widetilde{m}^{\min} \le m + \mathbf{D}u_\perp \le \widetilde{m}^{\max}\}\,. \tag{5.4}$$

To show that $\Sigma_\perp$ is non-empty, recall that the FVFT feasible set

$$\Sigma = \{u \in S_0^h \,|\, \widetilde{m}^{\min} \le m + \mathbf{D}u \le \widetilde{m}^{\max}\} \ne \emptyset.$$

Writing $u = u_0 + u_\perp$ and noting that $\mathbf{D}u = \mathbf{D}u_\perp$ implies that $\Sigma_\perp \ne \emptyset$ as well. Together with the strict convexity of the FVFT objective, it follows that Problem (a) has a unique minimizer. The following theorem summarizes our observations.

Theorem 5.2. *Let $u^{\mathsf{T}} = u_0^{\mathsf{T}} + u_\perp^{\mathsf{T}}$ be the orthogonal decomposition of the FVFT target. The optimal solution of (4.9) has the form*

$$u^{opt} = u_0^{\mathsf{T}} + u_\perp^{opt} \tag{5.5}$$

where $u_\perp^{opt}$ is the unique solution of Problem (a) in (5.3).

5.2. The structure of the FVMT optimal solution

We proceed to examine the hybrid FVMT formulation. The first result provides some additional information about the FVMT objective.

Proposition 5.3. *The FVMT objective is strictly convex on* $\ker \mathbf{D}^\perp$.

Proof. We need to show that for any $u, v \in \ker \mathbf{D}^\perp$, $u \ne v$ and $0 < \lambda < 1$, there holds

$$\|\lambda \mathbf{D}u + (1-\lambda)\mathbf{D}v\|_C^2 < \lambda\|\mathbf{D}u\|^2 + (1-\lambda)\|\mathbf{D}v\|_C^2\,.$$

Suppose this inequality does not hold. Then, there exists $0 < \lambda_0 < 1$ such that

$$\lambda_0^2\|\mathbf{D}u\|_C^2 + (1-\lambda_0)^2\|\mathbf{D}v\|^2 + 2\lambda_0(1-\lambda_0)(\mathbf{D}u, \mathbf{D}v)_C \ge \lambda_0\|\mathbf{D}u\|_C^2 + (1-\lambda_0)\|\mathbf{D}v\|_C^2\,,$$

which implies that

$$2(\mathbf{D}u, \mathbf{D}v)_C \ge \|\mathbf{D}u\|_C^2 + \|\mathbf{D}v\|_C^2 \quad \forall u, v \in \ker \mathbf{D}^\perp\,.$$

This is clearly a contradiction, which proves the assertion. ☐

Consider the restriction of the FVMT to the null-space complement:

$$\begin{cases} \underset{\widehat{u}_\perp \in \ker \mathbf{D}^\perp}{\text{minimize}} \quad \dfrac{1}{2}\|\mathbf{D}\widehat{u}_\perp - \mathbf{D}u^\mathsf{T}\|_C^2 \quad \text{subject to} \\ u_j^\mathsf{T} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x}) dV, \quad j = 1, \dots, S\,; \\ \widetilde{m}^{\min} \le m + \mathbf{D}\widehat{u}_\perp \le \widetilde{m}^{\max}\,. \end{cases} \tag{5.6}$$

The feasible set of (5.6) coincides with the set $\Sigma_\perp$ defined in (5.4). Therefore, the FVMT restriction has at least one non-trivial minimizer. On the other hand, Proposition 5.3 establishes the strict convexity of the FVMT objective on $\ker \mathbf{D}^\perp$. It follows that the QP (5.6) has a unique minimizer. With the help of these observations we can now characterize the class of optimal solutions of FVMT.

Theorem 5.4. *Any optimal solution of FVMT (4.18) has the form* $u^{opt} = u_0 + u_\perp^{opt}$ *where* $u_0 \in \ker \mathbf{D}$ *is arbitrary and* $u_\perp^{opt} \in \ker \mathbf{D}^\perp$ *is the unique solution of the FVMT restriction (5.6).*

Proof. Suppose that $u^{opt} \in \Sigma$ is an optimal solution of the FVMT with decomposition $u^{opt} = u_0^{opt} + u_\perp^{opt}$. Because $\mathbf{D}u^{opt} = \mathbf{D}u_\perp^{opt}$ it follows that $u_\perp^{opt} \in \Sigma_\perp$, while the optimality of u^{opt} implies that

$$\|\mathbf{D}u_\perp^{opt} - \mathbf{D}u^\mathsf{T}\|_C \le \|\mathbf{D}u_\perp - \mathbf{D}u^\mathsf{T}\|_C \quad \forall u_\perp \in \ker \mathbf{D}^\perp\,.$$

It follows that $u_\perp^{opt}$ solves (5.6). Because the latter has a unique minimizer, this proves the theorem. ☐

5.3. The relationship between FVFT and FVMT

While the FVFT and the FVMT solutions have similar structures, there are some important differences between them. The null-space component of the FVFT solution coincides with the null-space component u_0^T of the target, whereas the null-space component of the FVMT solution is an arbitrary element in $\ker \mathbf{D}$. As a result, the difference between these components can be arbitrarily large. On the other hand, Theorems 5.2 and 5.4 reveal that the solution components in $\ker \mathbf{D}^\perp$ solve similar strictly convex QPs, which have the same feasible set but different objectives. Therefore, it is not unreasonable to expect that these components are in some sense equivalent.

Let $u_\perp^F$ and $u_\perp^M$ denote the solution components of FVFT and FVMT in the null-space complement. Because $u_\perp^F \in \Sigma_\perp$ is the unique minimizer of Problem (a) in (5.3) and $u_\perp^M \in \Sigma_\perp$, the Poincare-Friedrichs inequality (3.8) implies that

$$\|u_\perp^F - u_\perp^\mathsf{T}\|_S \le \|u_\perp^M - u_\perp^\mathsf{T}\|_S \le K\|\mathbf{D}u_\perp^M - \mathbf{D}u_\perp^\mathsf{T}\|_C\,.$$

On the other hand, because $u_\perp^M \in \Sigma_\perp$ is the unique minimizer of the QP (5.6) and $u_\perp^F \in \Sigma_\perp$, we have that

$$\|\mathbf{D}u_\perp^M - \mathbf{D}u_\perp^\mathsf{T}\|_C \le \|\mathbf{D}u_\perp^F - \mathbf{D}u_\perp^\mathsf{T}\|_C \le C\|u_\perp^F - u_\perp^\mathsf{T}\|_S\,.$$

The following theorem provides a formal statement of this result.

Theorem 5.5. *The null-space components* $u_\perp^F$ *and* $u_\perp^M$ *of the FVFT and FVMT optimal solutions, respectively, are equivalent in the following sense: there exist constants* C_1 *and* C_2*, independent of the mesh size and such that*

$$C_1\|u_\perp^F - u_\perp^\mathsf{T}\|_S \le \|\mathbf{D}u_\perp^M - \mathbf{D}u_\perp^\mathsf{T}\|_C \le C_2\|u_\perp^F - u_\perp^\mathsf{T}\|_S\,. \tag{5.7}$$

5.4. The relationship between MVMT and FVFT/MT

The main goal of this section is to examine the relationship between the optimal solutions of the MVMT and the FVFT formulations. The intermediate FVMT formulation plays a crucial role in this task. Specifically, we first establish a connection between the optimal solutions of MVMT and FVMT. This connection and the result of Theorem 5.5 establish a notion of equivalence between MVMT and FVFT.

To distinguish the variables in the different OBR formulations we use the letter u for MVMT and the letter v for FVFT/MT. For convenience we restate the equivalence (4.17) of the MVMT and FVFT targets in this notation:

$$u^{\mathsf{T}} = \mathbf{D}v^{\mathsf{T}} .$$

Using this identity the QP (4.16) assumes the equivalent form

$$\begin{cases} \underset{\widehat{u}\in C^h}{\text{minimize}} \quad \dfrac{1}{2}\|\widehat{u} - \mathbf{D}v^{\mathsf{T}}\|_C^2 \qquad \text{subject to} \\ v_j^{\mathsf{T}} := \displaystyle\int_{r_j} \rho^h(\boldsymbol{x})dV, \quad j = 1, \dots, S\,; \\ \displaystyle\sum_{i=1}^{C} \widehat{u}_i = 0 \quad \text{and} \quad \widetilde{m}^{\min} \le m + \widehat{u} \le \widetilde{m}^{\max} . \end{cases} \tag{5.8}$$

We recall that the feasible set of (5.8) is given by

$$M = \{u \in C_0^h \,|\, \widetilde{m}^{\min} \le m + \widehat{u} \le \widetilde{m}^{\max}\} \neq \emptyset .$$

Let $v_{\perp}^{opt} \in \ker \mathbf{D}^{\perp}$ be the unique minimizer of the FVMT restriction (5.6) and define

$$u_F = \mathbf{D}v_{\perp}^{opt} .$$

Because the range of $\mathbf{D}$ is C_0^h and $v_{\perp}^{opt}$ is feasible for (5.6), i.e., $\widetilde{m}^{\min} \le m + \mathbf{D}v_{\perp}^{opt} \le \widetilde{m}^{\max}$, it follows that

$$u_F \in C_0^h \quad \text{and} \quad \widetilde{m}^{\min} \le m + u_F \le \widetilde{m}^{\max} .$$

In other words, u_F belongs to the feasible set M of the MVMT.

Consider now an arbitrary element $\widehat{u} \in M$. Because $\widehat{u} \in C_0^h$, the surjective property of $\mathbf{D}$ implies the existence of a $\widehat{v}_u \in \ker \mathbf{D}^{\perp}$, such that $\widehat{u} = \mathbf{D}\widehat{v}_u$. On the other hand, because $v_{\perp}^{opt}$ is the minimizer of (5.6), there holds

$$\|\widehat{u} - \mathbf{D}v_{\perp}^{opt}\|_C = \|\mathbf{D}\widehat{v}_u - \mathbf{D}v_{\perp}^{opt}\|_C \ge \|\mathbf{D}v_{\perp}^{opt} - \mathbf{D}v_{\perp}^{opt}\|_C = \|u_F - \mathbf{D}v_{\perp}^{opt}\|_C .$$

We have just established that u_F is feasible and

$$\|u_F - \mathbf{D}v_{\perp}^{opt}\|_C \le \|\widehat{u} - \mathbf{D}v_{\perp}^{opt}\|_C \quad \forall \widehat{u} \in M ,$$

that is u_F is an optimal solution of (5.8). The following theorem provides a formal statement of the equivalence between the MVMT and the FVMT solutions.

Theorem 5.6. *Let $u^{opt} \in M$ denote the unique minimizer of MVMT. There holds*

$$u^{opt} = \mathbf{D}v_{\perp}^{opt} \tag{5.9}$$

where $v_{\perp}^{opt}$ is the unique minimizer of the FVMT restriction (5.6).

Note that this result implies the equality of the MVMT and FVMT objective functions at the optimal solutions. The main result of this section is a direct consequence of this theorem and Theorem 5.5.

Corollary 5.7. *Let $u^{opt} \in M$ denote the unique minimizer of MVMT and $v_{\perp}^{F}$ the unique minimizer of Problem (a) in (5.3). Then, there exist constants C_1 and C_2, independent of the mesh size, such that*

$$C_1\|v_{\perp}^{F} - v_{\perp}^{\mathsf{T}}\|_S \le \|u^{opt} - \mathbf{D}v^{\mathsf{T}}\|_C \le C_2\|v_{\perp}^{F} - v_{\perp}^{\mathsf{T}}\|_S . \tag{5.10}$$

Because the null-space component of the FVFT solution is always equal to the null-space component of the target, this result shows that asymptotically MVMT and FVFT are equivalent, i.e., the accuracy of their solutions is comparable. The results in Section 7 confirm this.

6. Versatility of OBR

In this section we present three extensions of OBR that demonstrate the versatility of the divide-and-conquer optimization strategy. The first two examples demonstrate the flexibility afforded by the decoupling of the target definition from the enforcement of the physical properties. Section 6.1 introduces the notion of a reconstruction residual and defines *residual-dependent adaptable* targets that adjust to local solution properties. The example in Section 6.2 extends OBR to grids in spherical coordinates, which are common in many climate models. In both cases, the necessary modifications to OBR are confined to the target definition, leaving the rest of the formulation essentially unchanged. As a result, the extended formulations retain all theoretical and computational properties of the original OBR. In particular, the extensions do not require any changes in the design and implementation of the optimization algorithms, which are the core of the OBR approach. The third example extends OBR to a new class of conservative, feature-preserving optimization-based transport (OBT) algorithms. Because OBR provides the computational core of this extension, our new transport algorithms inherit the theoretical and computational properties of OBR. In particular, the developments in Sections 6.1–6.2 readily apply to the OBT algorithms, yielding adaptable, conservative, linearity and bounds-preserving transport schemes in $\mathbb{R}^d$ and on the sphere.

6.1. Optimization-based remap with adaptable targets

The targets in Section 4 were defined using a density reconstruction $\rho^h(\boldsymbol{x})$ that was assumed to be exact for linear functions. This approach works well for smooth densities but is less satisfactory for solutions with sharp features and discontinuities. In this section we take advantage of the separation between the target definition and the enforcement of the physical properties to develop targets that are better suited for piecewise smooth and discontinuous solutions.

Our main goal is to demonstrate the versatility of the optimization-based strategy rather than to present a full-fledged adaptable OBR, which is the focus of a forthcoming paper. Consequently, we consider a simple procedure that adapts the targets by using a *reconstruction residual* to modify the gradient of the reconstructed density $\rho^h(\boldsymbol{x})$.

To this end, we start with a piecewise density reconstruction $\rho^h(\boldsymbol{x})$ on the old mesh $C(\Omega)$ that is exact for linear densities and preserves the mean density on $C(\Omega)$:

$$\rho^h(\boldsymbol{x})|_{c_i} := \rho_i^h(\boldsymbol{x}) = \rho_i + \boldsymbol{g}_i \cdot (\boldsymbol{x} - \boldsymbol{b}_i)\,. \tag{6.1}$$

In (6.1) the symbol $\boldsymbol{g}_i$ denotes an approximation of $\nabla\rho$ on cell c_i. To define the reconstruction residual, note that the point value of any linear function at the barycenter of a domain $\mathscr{D}$ equals its mean value over this domain. Therefore, if the exact density is linear there holds

$$\rho(\boldsymbol{b}_i) = \frac{\int_{c_i} \rho(\boldsymbol{x}) dV}{\mu(c_i)} = \rho_i \quad \forall i = 1, \dots, C\,.$$

This property prompts the following definition of the reconstruction residual on cell c_i:

$$q_i = \sum_{j \in N(c_i)} |\rho_j - \rho_i^h(\boldsymbol{b}_j)|\,. \tag{6.2}$$

The vector q of all cell residuals belongs to the piecewise constant space C^h. Suppose that the exact density is linear. In this case there holds $\rho(\boldsymbol{b}_j) = \rho_j$ and since ρ^h is exact for linear functions we have that $\rho_i^h(\boldsymbol{b}_j) = \rho(\boldsymbol{b}_j) = \rho_j$, i.e., $\rho_j - \rho_i^h(\boldsymbol{b}_j) = 0$ for all $j \in N(c_i)$. It follows that $q_i = 0$, $\forall i \in C(\Omega)$, whenever the density is a globally linear function, thereby justifying the term "residual" for (6.2).

The value of q_i measures the deviation of the mean density values in the neighborhood of c_i from the values of a locally linear function, i.e., the reconstruction residual provides local information about the density behavior. In particular, if the value of q_i is large, this means that in the neighborhood of c_i the density significantly deviates from a linear function. We can use this information to locally adjust the reconstruction formula on all cells where q_i exceeds a certain threshold. To this end, consider a collection of real-valued functions α_i of a scalar argument such that

$$\alpha_i(\xi) \geq 1 \quad \text{and} \quad \alpha_i(0) = 1\,.$$

We define the adaptable reconstruction by multiplying the approximate gradient on each cell by a function of the reconstruction residual:

$$\rho^A(\boldsymbol{x})|_{c_i} := \rho_i^A(\boldsymbol{x}) = \rho_i + \alpha_i(q_i)\boldsymbol{g}_i \cdot (\boldsymbol{x} - \boldsymbol{b}_i)\,. \tag{6.3}$$

Substitution of ρ^h by ρ^A in (4.9), (4.16) and (4.18) defines the adaptable target extensions of these OBR formulations.

Clearly, (6.3) preserves the mean density values on the old mesh. Moreover, because the residual vanishes whenever ρ is globally linear, the adaptable reconstruction (6.3) remains exact for linear densities. As a result, the adaptable target extensions of the OBR formulations remain linearity preserving, yet their accuracy increases for solutions that differ significantly from linear functions. The examples in Section 7 confirm this conjecture. Finally, we note that since the targets are not required to be feasible, the use of (6.3) in no way affects the existence of optimal solutions in the OBR formulations.

It is instructive to compare the adaptable reconstruction (6.3) with a slope-limited reconstruction common in many transport methods:

$$\rho^L(\boldsymbol{x})|_{c_i} := \rho_i^L(\boldsymbol{x}) = \rho_i + \phi(R_i)\boldsymbol{g}_i \cdot (\boldsymbol{x} - \boldsymbol{b}_i)\,. \tag{6.4}$$

In (6.4) the number R_i is a slope ratio and $\phi(R_i)$ is a slope limiter [19]. The limiter in (6.4) carries the double burden of ensuring that the reconstruction is both monotone and as accurate as possible, with the former being the dominant goal. In contrast, the sole purpose of α_i in (6.3) is to improve the accuracy of the reconstruction without regard to whether or not the local bounds or monotonicity hold. A second significant difference is that the slope limiter depends on slope ratios, i.e., it is not a residual-based quantity. As a result, $\phi(R)$ usually exhibits a "single behavior mode" that can sometimes change solution features. For example, the SuperBee [40] and the Barth-Jespersen [41] limiters can turn smooth waves into square waves, and the former tends to steepen gradients [19, 46].

6.2. Extension to spherical geometry

As a second demonstration of the versatility of our approach, we extend the OBR to grids in spherical coordinates, which are commonly used in climate models. This extension relies on simple modifications of the target computation but does not in any way affect the optimization core of the OBR, nor does it require changes in the definition of the constraints.

We consider a geophysical coordinate system with latitude $\theta \in [-\pi/2, \pi/2]$ and longitude $\lambda \in [0, 2\pi]$ on a sphere with unit radius. Away from the poles, the extension of OBR to spherical coordinates boils down to the proper evaluation of the entities and quantities entering the linear density reconstruction and its subsequent integration on the swept regions in spherical coordinates.

As in the case of $\mathbb{R}^d$, the unsigned measure of a set $\mathscr{D}$ on the surface of the sphere is

$$\mu(\mathscr{D}) = \int_{\mathscr{D}} dV\,,$$

whereas the application of the Green's theorem

$$\int_{\mathscr{D}} dV = -\int_{\partial\mathscr{D}} \sin\theta d\lambda\,.$$

in conjunction with an orientation choice for $\partial\mathscr{D}$ defines the signed measure $\mu^\star(\mathscr{D})$.

In the case of spherical coordinates the barycenter formula for a cell c_j with position vector $\boldsymbol{s}$ on the surface of the sphere

$$\boldsymbol{b}_j = \frac{\displaystyle\int_{c_j} \boldsymbol{s}\, dV}{\mu(c_j)} \,,$$

assumes the component-wise form

$$\boldsymbol{b}_{\theta,j} = \frac{\displaystyle\int_{c_j} \theta\, dV}{\mu(c_j)} \,, \quad \text{and} \quad \boldsymbol{b}_{\lambda,j} = \frac{\displaystyle\int_{c_j} \lambda \cos\theta\, dV}{\mu(c_j)} \,.$$

To compute the barycenter components we follow the approach in [42] and use the Green's theorem to convert the integral in the numerator into a boundary integral:

$$\boldsymbol{b}_{\theta,j} = \frac{\displaystyle -\int_{\partial c_j} (\cos\theta + \theta\sin\theta)\, d\lambda}{\mu(c_j)} \,, \quad \text{and} \quad \boldsymbol{b}_{\lambda,j} = \frac{\displaystyle -\int_{\partial c_j} \lambda/2(\sin\theta\cos\theta + \theta)\, d\lambda}{\mu(c_j)} \,.$$

The formula for a linear density reconstruction that preserves the mean density (ρ_i) is the same in spherical coordinates as in $\mathbb{R}^d$:

$$\rho^h(\boldsymbol{s})|_{c_i} := \rho^h_i(\boldsymbol{s}) = \rho_i + \boldsymbol{g}_i \cdot (\boldsymbol{s} - \boldsymbol{b}_i) \,,$$

except that the vector $\boldsymbol{g}_i$ approximates the gradient of ρ in spherical coordinates:

$$\boldsymbol{g}_i = (g^\theta_i, g^\lambda_i) \approx \left(\frac{\partial\rho}{\partial\theta}, \frac{1}{\cos\theta}\frac{\partial\rho}{\partial\lambda}\right)\Big|_{c_i} \,.$$

Thus, in latitude and longitude components the mean-preserving reconstruction has the form

$$\rho^h(\boldsymbol{s})|_{c_i} = \rho_i + \left(g^\theta_i\right)(\theta - \boldsymbol{b}_{\theta,i}) + \left(g^\lambda_i\right)(\lambda\cos\theta - \boldsymbol{b}_{\lambda,i}) \,.$$

Once the density reconstruction in spherical coordinates is available, the rest of the OBR algorithm functions exactly as in the $\mathbb{R}^d$ case. In particular, we have the FVFT and the MVMT versions of OBR, as well as their adaptable target extensions in spherical coordinates. The latter follows verbatim the process outlined in Section 6.1.

6.3. Extension to optimization-based transport

This section demonstrates another aspect of the versatility of the OBR approach. Specifically, we combine the optimization framework developed in Sections 2–4 with the incremental remapping approach [27] to obtain a new class of feature-preserving methods for the scalar transport equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\rho\boldsymbol{v} = 0 \;\text{ on } \Omega\times[0,T] \quad\text{and}\quad \rho(\boldsymbol{x},0) = \rho^0(\boldsymbol{x}) \,. \tag{6.5}$$

Here $T > 0$ is the final time, $\rho(\boldsymbol{x},t)$ is a positive density function on $\Omega\times[0,T]$ with the initial distribution $\rho^0(\boldsymbol{x})$, and $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x},t)$ is a given velocity field.

Let $\Omega(t)$ denote the deformed (Lagrangian) domain at time $t \in [0,T]$, i.e., the domain whose points move along the characteristic lines $d\boldsymbol{x}/dt = \boldsymbol{v}$. We assume that $\boldsymbol{v}$ is such that $\Omega(t_1) = \Omega(t_2)$ for any $t_1, t_2 \in [0,T]$. This assumption enables the fulfillment of the restrictions on the positions of the boundary vertices in Section 3.1, but is not otherwise critical. Let $V(t) \subset \Omega(t)$ be an arbitrary Lagrangian control volume with mass

$$m_V(t) = \int_{V(t)} \rho(\boldsymbol{x},t) dV \,.$$

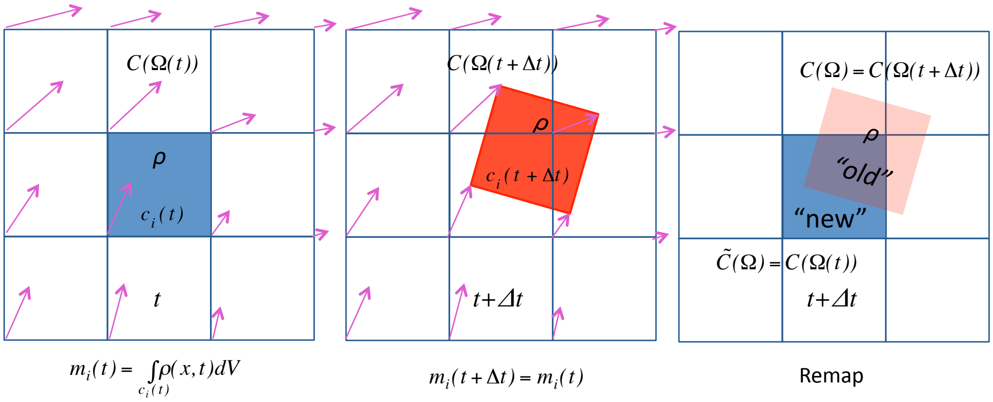


Figure 3: A single step of a transport algorithm based on forward incremental remapping. The left pane shows the departure grid $C(\Omega(t))$. The middle pane illustrates the Lagrangian step in which the blue cell from the departure grid evolves into the red cell on the arrival grid $C(\Omega(t+\Delta t))$. The right pane shows the remap of mass and density from the arrival ("old") grid to the departure ("new") grid.

Equation (6.5) is equivalent to

$$\frac{d}{dt} m_V(t) = 0 \quad \forall V(t) \subset \Omega(t)\,. \tag{6.6}$$

The conservation of mass in Lagrangian volumes (6.6) is the basis for the incremental remapping approach [27] for the solution of (6.5). Note that setting $V(t) = \Omega(t)$ in (6.6) yields the conservation of total mass.

Let $C(\Omega(t))$ denote a current or "departure" grid on the deformed domain $\Omega(t)$ at the current time t. The projection of the cells $c_i(t)$ of the departure grid into a future time $t + \Delta t$ yields a future or "arrival" grid $C(\Omega(t+\Delta t))$ with cells $c_i(t+\Delta t)$. Assuming that the time step is small enough, the arrival grid is a valid partition of $\Omega(t+\Delta t)$, which satisfies the assumptions in Section 3.1. Owing to (6.6) the mass in each cell is conserved,

$$m_i^{ex}(t+\Delta t) = \int_{c_i(t+\Delta t)} \rho(\boldsymbol{x}, t+\Delta t)\, dV = \int_{c_i(t)} \rho(\boldsymbol{x}, t)\, dV = m_i^{ex}(t)\,, \tag{6.7}$$

and the mean cell density on the arrival grid is

$$\rho_i^{ex}(t+\Delta t) = \frac{m_i^{ex}(t+\Delta t)}{\mu(c_i(t+\Delta t))} = \frac{m_i^{ex}(t)}{\mu(c_i(t+\Delta t))}\,. \tag{6.8}$$

Suppose that $\rho(t) \in C^h$ and $m(t) \in C^h$ approximate $\rho^{ex}(t)$ and $m^{ex}(t)$ on the departure grid $C(\Omega(t))$. From (6.7)–(6.8) it follows that

$$m_i(t+\Delta t) = m_i(t) \quad \text{and} \quad \rho_i(t+\Delta t) = \frac{m_i(t+\Delta t)}{\mu(c_i(t+\Delta t))} = \frac{m_i(t)}{\mu(c_i(t+\Delta t))}$$

approximate the cell masses and the mean cell density values, respectively, on the arrival grid $C(\Omega(t+\Delta t))$. The forward incremental remapping approach treats the departure and the arrival grids as the "new" and "old" grids respectively, i.e., we set $\widetilde{C}(\Omega) = C(\Omega(t))$ and $C(\Omega) = C(\Omega(t+\Delta t))$. Remapping $\rho(t+\Delta t)$ and $m(t+\Delta t)$ from the arrival grid to the departure grid yields approximations $\widetilde{\rho}$ and $\widetilde{m}$ of the mean cell density and the cell masses at the future time $t + \Delta t$ *but with respect to the departure grid cells*:

$$\widetilde{m}_i \approx \widetilde{m}_i^{ex} = \int_{c_i(t)} \rho(\boldsymbol{x}, t+\Delta t)\, dV\,; \quad \widetilde{\rho}_i \approx \widetilde{\rho}^{ex} = \frac{\int_{c_i(t)} \rho(\boldsymbol{x}, t+\Delta t)\, dV}{\mu(c_i(t))} = \frac{\widetilde{m}_i^{ex}}{\mu(c_i(t))}\,.$$

In other words, repeatedly projecting the same fixed departure grid forward in time and then immediately remapping the mass and the density back to it is equivalent to finding an approximate solution of (6.5) on the fixed departure grid; see Figure 3. This approach is also known as conservative semi-Lagrangian transport and is commonly used in climate models [29, 28, 43]. By combining the forward incremental remapping with an optimization-based remap we obtain a new class of optimization-based, feature preserving transport algorithms.

In summary, assuming that we are given the approximations $m(t)$ and $\rho(t)$ on some fixed departure grid $\widetilde{C}(\Omega) = C(\Omega(t))$, one step of the optimization-based transport method comprises the following stages:

- Project departure grid to arrival grid: $C(\Omega(t)) \mapsto C(\Omega(t+\Delta t))$;
- Lagrangian transport: $m_i(t+\Delta t) = m_i(t)$ and $\rho_i(t+\Delta t) = m_i(t)/\mu(c_i(t+\Delta t))$, for $i = 1, \dots C$;
- Optimization-based remap: $m(t+\Delta t) \mapsto \widetilde{m}$ and $\rho(t+\Delta t) \mapsto \widetilde{\rho}$, for $i = 1, \dots C$.

The optimization-based transport (OBT) is a conservative semi-Lagrangian scheme. Because OBR provides the computational core of this scheme, OBT inherits all theoretical and computational properties of the algorithms in Section 4. Furthermore, by using the adaptable targets in Section 6.1 we readily obtain an adaptable version of OBT, whereas the extension of OBR to spherical geometry provides an adaptable version of OBT for transport on the sphere.

7. Numerical studies

The numerical behavior of optimization algorithms designed for an efficient solution of the FVFT and MVMT formulations is investigated in the context of mesh remapping in [2], where they are also compared to the explicit FCR scheme. This study shows that for small mesh displacements the qualitative ("eyeball") and quantitative (asymptotic) accuracies of FVFT, MVMT and FCR, respectively, are very similar. The study also shows that FVFT and MVMT are considerably more robust and more accurate than FCR when the mesh displacements are large. The application of FVFT and FCR to scalar transport is studied in [44], with similar conclusions. However, [44] also reveals that the computational cost of FVFT, while exhibiting linear complexity with respect to the mesh size, can be $\mathscr{O}(10)$ larger than the computational cost of FCR in transport applications. This renders FVFT-based transport impractical in most cases.

The numerical studies presented here focus on answering two questions:

- Does the simpler structure of the MVMT optimization problem yield an algorithm that is competitive with the explicit FCR scheme in the context of transport applications?
- Does the flexibility of optimization-based remap algorithms to use adaptable high-order targets yield qualitative improvements in scalar transport?

It turns out that the answers to both questions are positive.

7.1. Transport in a plane

Our numerical studies begin with a comparison of the computational cost and numerical accuracy of FCR, FVFT, MVMT and MVMT-a; by MVMT-a we denote the MVMT formulation (4.16) with the adaptable target

$$\rho_i^A(\boldsymbol{x}) = \rho_i + \alpha_i(q_i)\boldsymbol{g}_i \cdot (\boldsymbol{x} - \boldsymbol{b}_i) \quad \forall i = 1, \dots, C\,,$$

where for the given constants $\gamma_1, \gamma_2 > 0$,

$$\alpha_i(q_i) = \begin{cases} 1 & \text{if } q_i / \max\limits_{i=1,\dots,C}\{q_i\} \le \gamma_1 \\ 1 + \gamma_2\, q_i / \max\limits_{i=1,\dots,C}\{q_i\} & \text{otherwise}\,. \end{cases}$$

The FVFT optimization algorithm is described in [45]. The simpler MVMT algorithm is developed in [2].

The first example is a rotation of a solid body given by the combination of a smooth hump, a cone and a slotted cylinder. The setup is identical to that described in [45] and goes back to LeVeque [46]. The

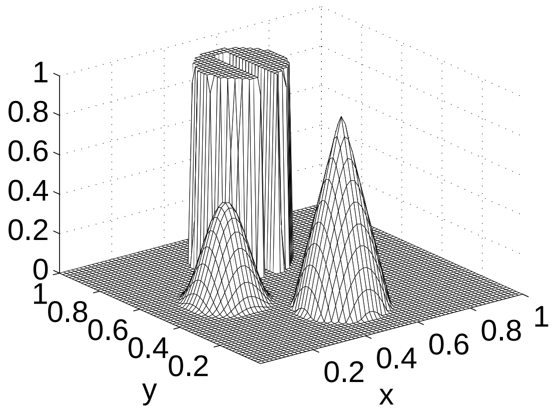


Figure 4: Initial data for the solid body rotation test.

initial condition is depicted in Figure 4, over the domain $\Omega = [0,1]\times[0,1]$. The smooth hump is given by $\rho(x_1,x_2,0) = 0.25(1+\cos(\pi r(x_1,x_2)))$, where $r(x_1,x_2) = \min\{\sqrt{(x_1-0.25)^2+(x_2-0.5)^2}, 0.15\}/0.15$. We use a cone and a slotted cylinder of radius 0.15 and height 1, centered at (0.5,0.25) and (0.5,0.75), respectively. The slot in the cylinder is created by removing the region [0.475,0.525]×[0.6,0.85]. The velocity field is given by $\boldsymbol{v}(x_1,x_2) = (-x_2+0.5,\, x_1-0.5)$. For time discretization we use RK4, the fourth-order Runge-Kutta method, in all numerical experiments.

The results of a full rotation on a 128×128 mesh, with the final time $T = 2\pi$ and 810 time steps, are given in Figure 5 for each of the methods. Here we use the MVMT-a parameters $\gamma_1 = 0.2$ and $\gamma_2 = 0.3$. The figure demonstrates that FCR, FVFT and MVMT perform similarly in the eyeball norm, whereas MVMT-a stands out as being qualitatively more accurate. This is confirmed in Figure 6, where it is shown that MVMT-a resolves the back side and the slot of the cylinder better than FCR, while the cone is recovered equally well by the methods.

All methods perform similarly when it comes to asymptotic accuracy, see Table 1. However, as expected from the qualitative results, MVMT-a shows the lowest absolute errors. Table 2 confirms that FVFT is the least efficient method, trailing behind FCR by a factor of $\mathscr{O}(10)$. This is in agreement with our findings in [44], and due to the large, globally coupled linear systems that must be solved at every optimization iteration. On the other hand, the computational cost of MVMT is marginally *lower* than the cost of FCR, while MVMT-a is only slightly less efficient. The huge improvement in performance over FVFT can be traced back to a mostly decoupled, cell-local optimization algorithm, with the coupling occurring in a single constraint that is easily satisfied, see [2].

Solid-body rotation (L_1 error)

		FCR		**FVFT**		**MVMT**		**MVMT-a**	
# cells	# steps	L_1 error	rate	L_1 error	rate	L_1 error	rate	L_1 error	rate
64×64	408	3.42e-2	—	3.59e-2	—	3.93e-2	—	3.13e-2	—
128×128	810	1.94e-2	0.82	2.05e-2	0.80	2.34e-2	0.81	1.73e-2	0.86
256×256	1614	1.12e-2	0.81	1.19e-2	0.80	1.39e-2	0.80	1.12e-2	0.75

Table 1: Comparison of the L_1 errors of FCR, FVFT, MVMT and MVMT-a with respect to the initial condition given in Figure 4. The rates of convergence are similar for all methods; MVMT-a shows the lowest absolute errors.

Our next test is a left-to-right-and-back translation of the slotted cylinder, see Figure 7. For this experiment, the slotted cylinder is centered at (0.5,0.5). The final time is $T = 6$ and the velocity field is given by $\boldsymbol{v}(x_1,x_2,t) = (1,0)$ for $t \in [0,3)$ and $\boldsymbol{v}(x_1,x_2,t) = (-1,0)$ for $t \in [3,6]$. We note that our implementation utilizes periodic boundaries for the computational domain Ω and that the slotted cylinder

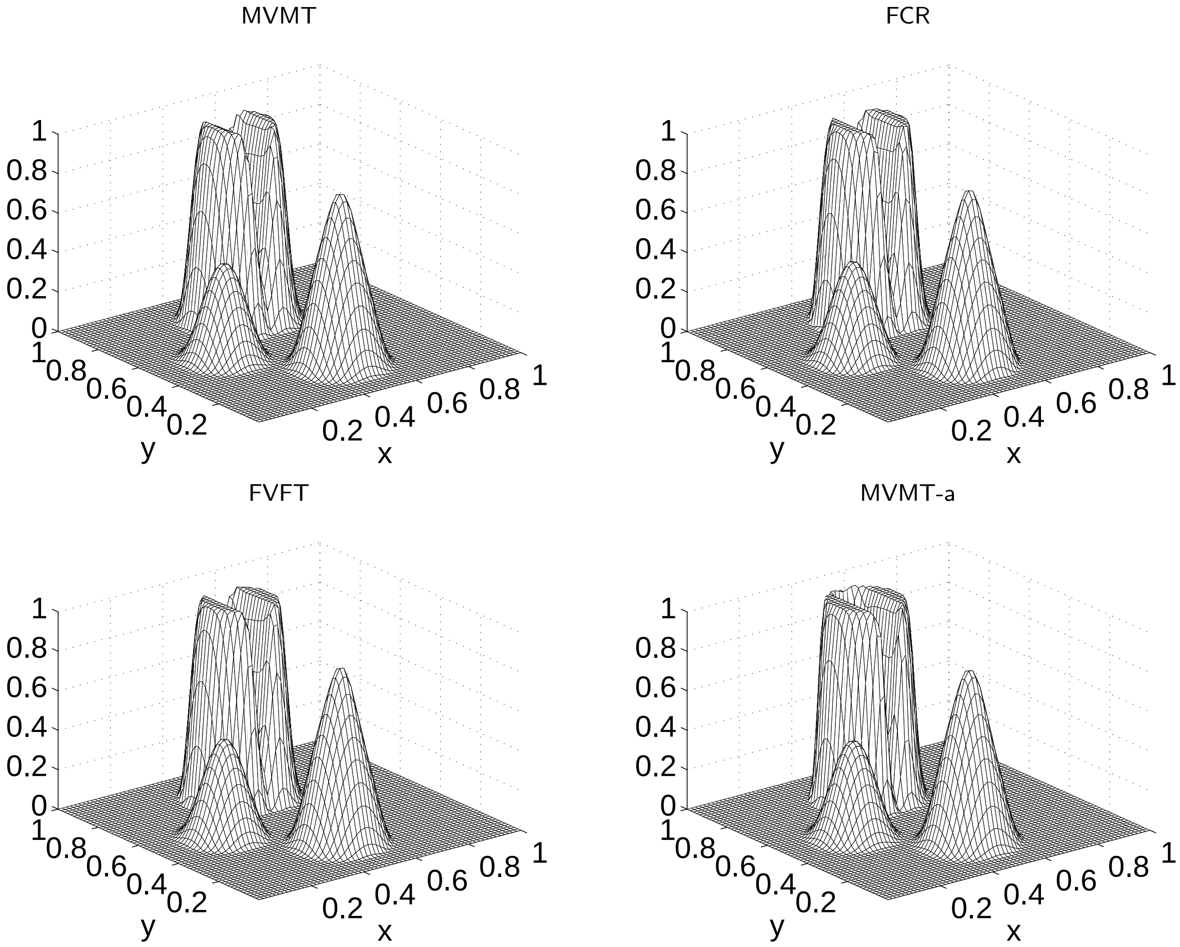


Figure 5: Transport results for the solid-body rotation test, for one full revolution (810 time steps) on a 128×128 mesh; FCR, FVFT and MVMT perform similarly in the eyeball norm, whereas MVMT-a stands out as being qualitatively more accurate.

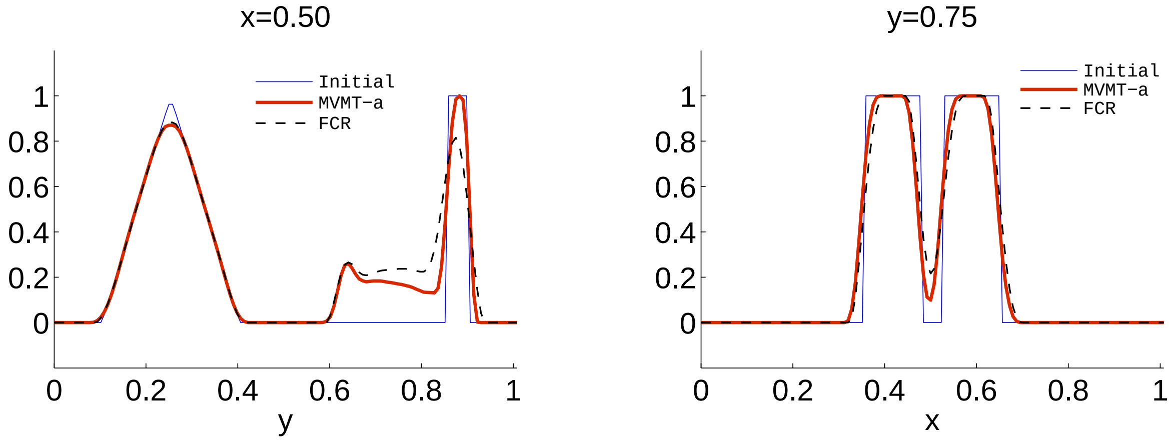


Figure 6: Comparison of the qualitative accuracy of MVMT-a and FCR for the solid-body rotation test. Left pane: $x = 0.50$ slice. Right pane: $y = 0.75$ slice. The MVMT-a scheme resolves the back side and the slot of the cylinder better than FCR.

Solid-body rotation (timings)

		FCR	**FVFT**		**MVMT**		**MVMT-a**	
# cells	# steps	time(sec)	time(sec)	ratio	time(sec)	ratio	time(sec)	ratio
64×64	408	3.3	63.7	19.3	3.4	1.0	3.8	1.1
128×128	810	26.4	496.4	18.8	26.2	1.0	28.8	1.1
256×256	1614	229.1	3464.2	15.1	222.7	1.0	230.9	1.0

Table 2: Comparison of the computational costs of FCR, FVFT, MVMT and MVMT-a as measured by Matlab^{TM} wall-clock times in seconds, on a single Intel Xeon X5680 3.33GHz processor, for the solid-body (smooth hump, cone, slotted cylinder) rotation test. Ratios of run times of FVFT, MVMT and MVMT-a with respect to FCR are included. The costs of MVMT and MVMT-a are almost identical to the cost of FCR, while FVFT is 15 to 20 times slower.

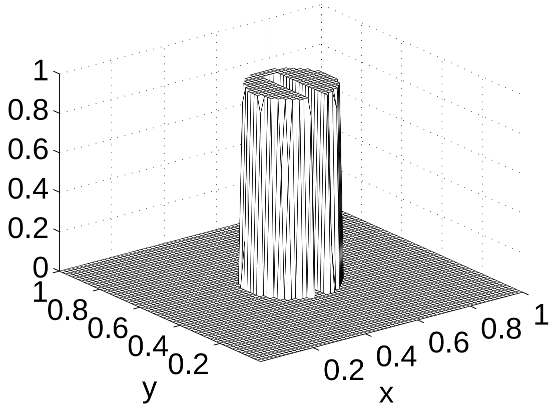


Figure 7: Initial data for the solid body translation test.

traverses the domain six times during the experiment (three times to the right and three times to the left).

The results of the solid-body translation on a 128×128 mesh, with a total of 1092 time steps, are given in Figure 8 for FCR and MVMT-a. Here we use the MVMT-a parameters $\gamma_1 = 0.1$ and $\gamma_2 = 0.75$. In this experiment MVMT-a outperforms FCR significantly in the eyeball norm, fully preserving the slot in the cylinder. Figure 9 confirms this result. Table 3 shows that with respect to the initial condition given in Figure 7 MVMT-a exhibits faster convergence and lower absolute errors.

Solid-body translation (timings and L_1 error)

		FCR	**MVMT-a**		**FCR**		**MVMT-a**	
# cells	# steps	time(sec)	time(sec)	ratio	L_1 error	rate	L_1 error	rate
64×64	546	4.2	4.9	1.2	2.77e-2	—	2.75e-2	—
128×128	1092	35.2	37.8	1.1	1.59e-2	0.80	1.34e-2	1.04
256×256	2178	282.2	295.5	1.0	9.45e-3	0.78	6.70e-3	1.02

Table 3: (1) Comparison of the computational costs of FCR and MVMT-a as measured by Matlab^{TM} wall-clock times in seconds, on a single Intel Xeon X5680 3.33GHz processor, for the slotted-cylinder translation test. Ratios of run times of MVMT-a with respect to FCR are included. The cost of MVMT-a is very close to the cost of FCR. (2) Comparison of the L_1 errors with respect to the initial condition given in Figure 7. MVMT-a exhibits lower absolute errors and faster convergence.

7.2. Transport on a sphere

Two numerical studies are performed on the spherical domain to compare the computational cost and numerical accuracy of FCR and MVMT-a. The first computational example on the sphere uses a non-

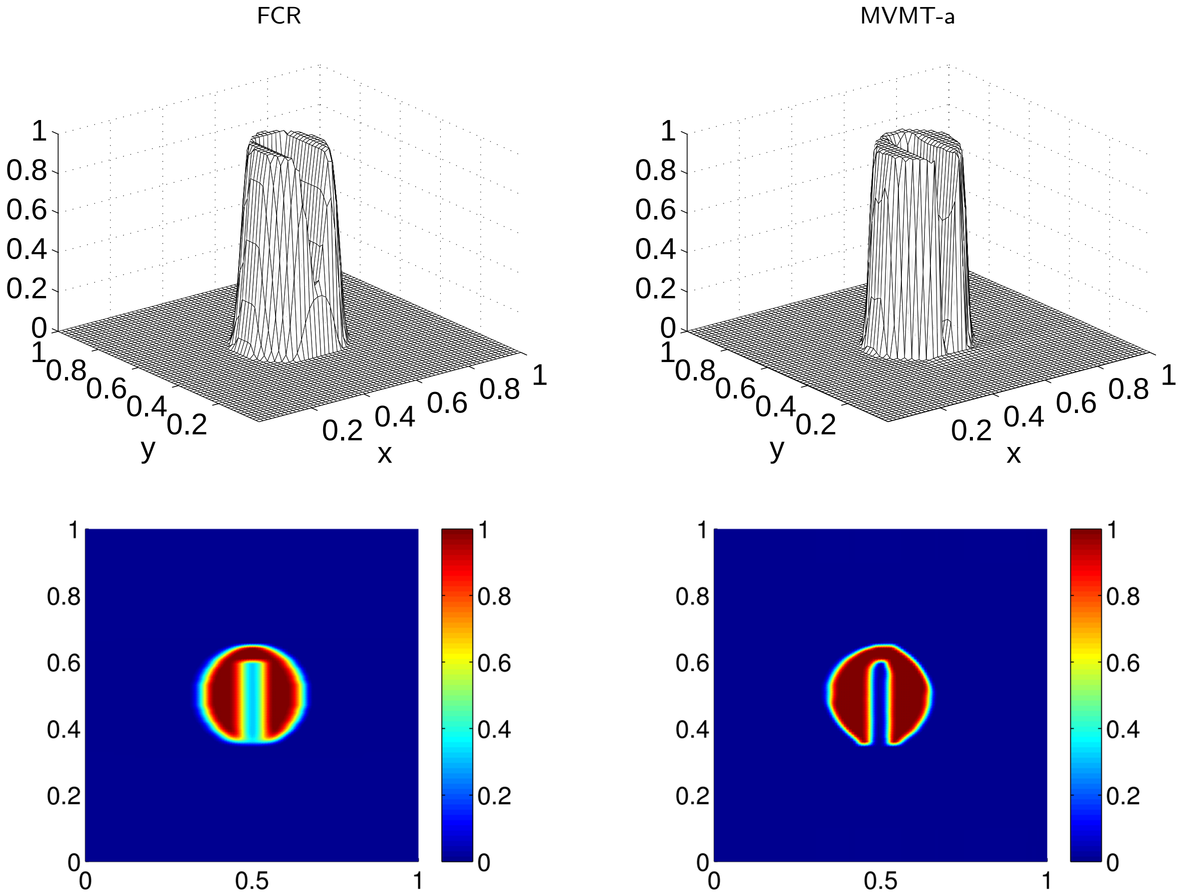


Figure 8: Transport results for the solid-body translation test, for 546 time steps to the right and 546 time steps to the left on a 128×128 mesh. In this experiment MVMT-a fully preserves the slot in the cylinder, while FCR is significantly more diffusive.

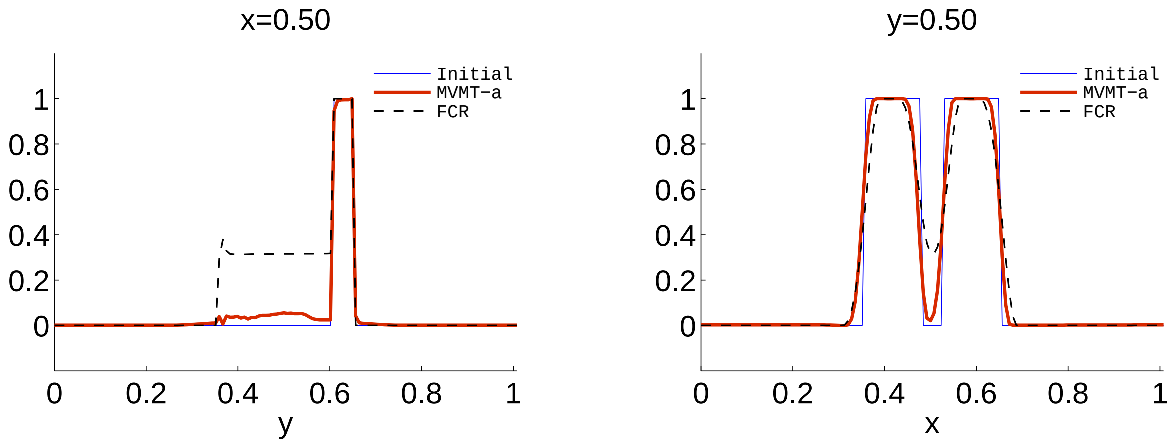


Figure 9: Comparison of the qualitative accuracy of MVMT-a and FCR for the solid-body translation test. Left pane: $x = 0.50$ slice. Right pane: $y = 0.50$ slice. MVMT-a resolves the back side and the slot of the cylinder much better than FCR.

divergent deformational velocity field due to Nair and Lauritzen [47]

$$u(\lambda, \theta, t) = 2\sin^2(\lambda)\sin(2\theta)\cos(\pi t/T)$$
$$v(\lambda, \theta, t) = 2\sin(2\lambda)\cos(\theta)\cos(\pi t/T)$$

with the period T set to 5. An initial density distribution consisting of two notched cylinders with radii $r = 1/2$, height $h = 1$, and initial positions $(\lambda_0, \theta_0) = (5\pi/6, 0)$ and $(\lambda_1, \theta_1) = (7\pi/6, 0)$ is used. The two slotted cylinders are oriented symmetrically with respect to the deformational flow field in this configuration. In spherical coordinates the notched cylinder implementation requires the computation of the great circle distance between an arbitrary point (λ, θ) and a cylinder center (λ_i, θ_i), which is defined as

$$d_i(\lambda, \theta) = \arccos\left(\sin\theta_i \sin\theta + \cos\theta_i \cos\theta \cos(\lambda - \lambda_i)\right).$$

With this definition the initial configuration of the notched cylinders, shown in Figure 10, may be expressed in spherical coordinates as

$$\rho(\lambda, \theta) = \begin{cases} h & \text{if } d_i < r \text{ and } |\lambda - \lambda_i| \geq r/6 \text{ for } i = 1, 2 \\ h & \text{if } d_0 < r \text{ and } |\lambda - \lambda_0| < r/6 \text{ and } \theta - \theta_0 < -5r/12 \\ h & \text{if } d_1 < r \text{ and } |\lambda - \lambda_1| < r/6 \text{ and } \theta - \theta_1 > 5r/12 \\ 0 & \text{otherwise.} \end{cases}$$

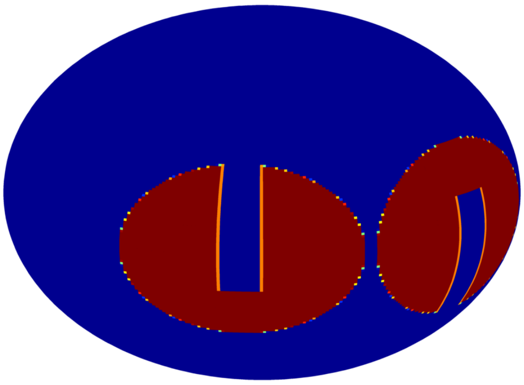


Figure 10: Initial data for the transport tests on the sphere.

The test case is run on three meshes with resolutions of 3°, 1.5°, and 0.75°, using RK4 for the time discretization. Results at a final time $T = 5$ for the 0.75° grid are shown in Figure 11. The MVMT-a result, where parameters $\gamma_1 = 0.1$ and $\gamma_2 = 0.75$ were used, appears qualitatively better than the FCR result particularly for the cylinder on the right. As seen in Table 4 both methods have similar rates of convergence, but MVMT-a has smaller absolute errors than FCR and the computational costs of the methods are nearly the same.

Deformational velocity field on the sphere (timings and L_1 error)

mesh	# steps	FCR time(sec)	MVMT-a time(sec)	ratio	FCR L_1 error	rate	MVMT-a L_1 error	rate
3°	600	23.0	24.2	1.1	4.34e-2	—	3.60e-2	—
1.5°	1200	187.7	190.0	1.0	2.85e-2	0.61	2.27e-2	0.66
0.75°	2400	1644.4	1717.7	1.0	1.67e-2	0.69	1.40e-2	0.68

Table 4: (1) Comparison of the computational costs of FCR and MVMT-a as measured by Matlab^{TM} wall-clock times in seconds, on a single Intel Xeon X5680 3.33GHz processor, for the two slotted-cylinder deformation test on the sphere. Ratios of run times of MVMT-a with respect to FCR are included. The cost of MVMT-a is very close to the cost of FCR. (2) Comparison of the L_1 errors with respect to the initial condition given in Figure 10. MVMT-a exhibits lower absolute errors and comparable convergence.

The second computational example is a solid body rotation test using the temporally constant zonal flow field

$$\begin{aligned} u(\lambda, \theta) &= 2\pi\left(\cos(\theta)\cos(\alpha) + \cos(\lambda)\sin(\theta)\sin(\alpha)\right) \\ v(\lambda, \theta) &= 2\pi \sin(\lambda)\sin(\alpha), \end{aligned}$$

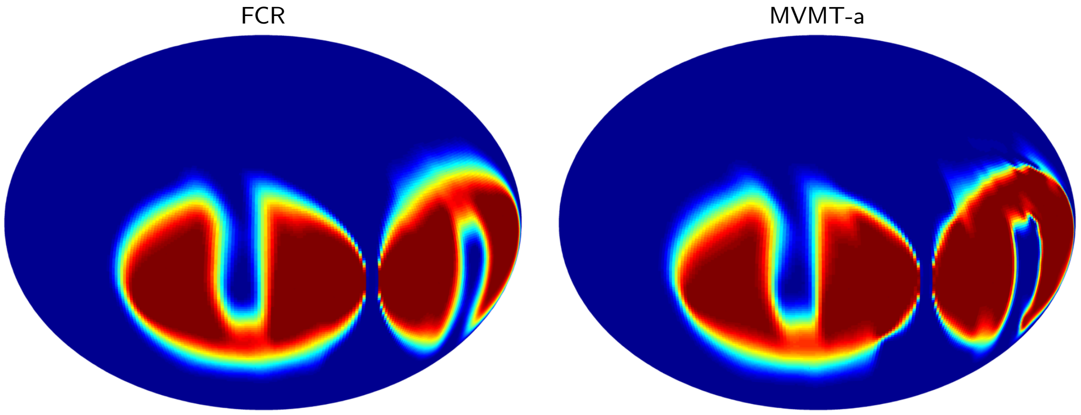


Figure 11: Transport results for the deformational flow test on the sphere at a final time $T = 5$ after 2400 time steps on a 0.75° mesh. The MVMT-a results appear sharper than the FCR results, particularly for the cylinder on the right.

where the rotation angle α provides the orientation of the flow [29]. In the examples shown here α is set to zero resulting in a flow oriented along the equator. The initial density configuration used in the previous experiment and shown in Figure 10 is also used in this case. The notched cylinders are rotated around the domain twice, first in one direction and then in the opposite direction using RK4 for the time discretization. In this experiment MVMT-a with parameters $\gamma_1 = 0.1$ and $\gamma_2 = 0.75$ exhibits a higher rate of convergence and lower absolute errors than FCR, as seen in Table 5, while the computational cost of MVMT-a is only slightly higher than that of FCR. Results for two full rotations of the cylinders around the equator on the 0.75° mesh with final time $T = 2$ for 1920 time steps are shown in Figure 12 where the improved performance of MVMT-a over FCR can be seen qualitatively.

Solid-body translation on the sphere (timings and L_1 error)

		FCR	**MVMT-a**		**FCR**		**MVMT-a**	
mesh	# steps	time(sec)	time(sec)	ratio	L_1 error	rate	L_1 error	rate
3°	480	17.4	18.2	1.0	3.25e-2	—	2.79e-2	—
1.5°	960	132.5	151.6	1.1	1.99e-2	0.78	1.36e-3	1.04
0.75°	1920	1184.5	1379.0	1.2	1.10e-2	0.78	5.41e-3	1.18

Table 5: (1) Comparison of the computational costs of FCR and MVMT-a as measured by Matlab^{TM} wall-clock times in seconds, on a single Intel Xeon X5680 3.33GHz processor, for the slotted-cylinder translation test on the sphere. Ratios of run times of MVMT-a with respect to FCR are included. The cost of MVMT-a is very close to the cost of FCR. (2) Comparison of the L_1 errors with respect to the initial condition given in Figure 10. MVMT-a exhibits lower absolute errors and faster convergence.

8. Conclusions

In this paper we investigated a divide-and-conquer optimization-based strategy for the formulation of *feature-preserving* numerical methods. We applied this strategy to the constrained interpolation (remap) of a scalar conserved quantity and derived a collection of optimization-based remap (OBR) formulations.

We presented and studied three instances of OBR formulations, which differ by the definition of their targets, constraint sets and objective functions. While mathematically all three yield equivalent solutions, the mass variable mass target (MVMT) formulation proved to be the most efficient algorithmically, with the cost essentially equivalent to that of the explicit Flux Corrected Remap (FCR) approach.

The use of adaptable targets and the extensions to spherical geometry and to the conservative, bounds-preserving optimization-based transport clearly demonstrate the versatility of our approach.

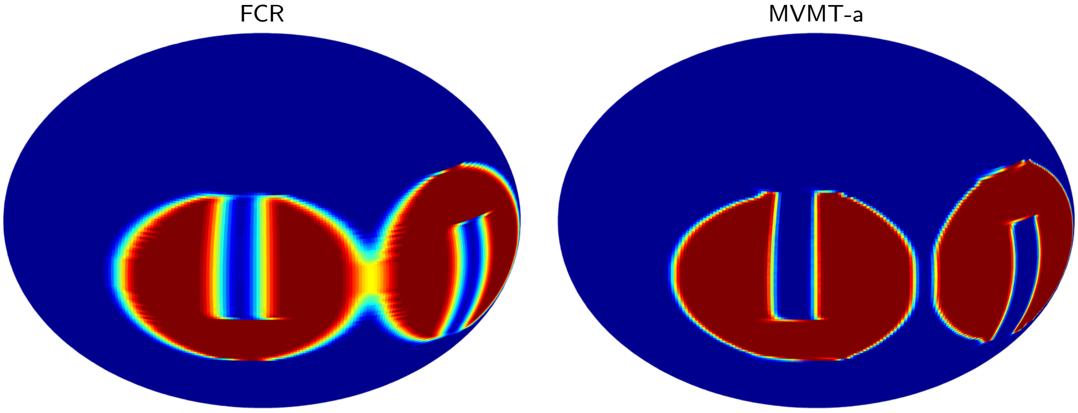


Figure 12: Transport results for the solid-body rotation test on the sphere, for two revolutions (1920 time steps) on a 0.75° mesh. The MVMT-a results provide a qualitatively better match to the initial conditions than the FCR results.

Acknowledgments

All authors acknowledge funding by the DOE Office of Science Advanced Scientific Computing Research (ASCR) Program. DR and KP also acknowledge funding by the Advanced Simulation & Computing (ASC) Program. Special thanks go to M. Shashkov who introduced us to the remap problem, guided us through its subtleties and prompted our initial interest in the use of optimization in this context. Our colleague J. Young helped with the formulation and implementation of some of the optimization algorithms used in the approach. Technical discussions and informal conversations with R. Abgrall, A. Barlow, E. Cyr, D. Kuzmin, A. Robinson, J. Shadid, G. Scovazzi, and many other colleagues provided useful insights and perspectives.

References

[1] P. Bochev, D. Ridzal, G. Scovazzi, M. Shashkov, Formulation, analysis and numerical study of an optimization-based conservative interpolation (remap) of scalar fields for arbitrary lagrangian-eulerian methods, Journal of Computational Physics 230 (2011) 5199 – 5225.

[2] P. Bochev, D. Ridzal, M. Shashkov, Fast optimization-based conservative remap of scalar fields through aggregate mass transfer, Journal of Computational Physics Submitted (2012).

[3] D. N. Arnold, Differential complexes and numerical stability, in: L. Tatsien (Ed.), Proceedings of the International Congress of Mathematicians, Beijing, volume I, pp. 137–157.

[4] D. N. Arnold, P. Bochev, R. Lehoucq, R. Nicolaides, M. Shashkov (Eds.), Compatible Discretizations. Proceedings of IMA Hot Topics workshop on Compatible discretizations, IMA Volume 142, Springer Verlag, 2006.

[5] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, Acta Numerica 15 (2006) 1–155.

[6] A. Dezin, Multidimensional Analysis and Discrete Models, CRC Presss, Boca Raton, FL, 1995.

[7] M. Shashkov, Conservative Finite Difference Methods on General Grids, CRC Press, Boca Raton, FL, 1995.

[8] P. G. Ciarlet, P. A. Raviart, Maximum principle and uniform convergence for the finite element method, Computer Methods in Applied Mechanics and Engineering 2 (1973) 17 – 31.

[9] A. Draganescu, T. F. Dupont, L. R. Scott, Failure of the discrete maximum principle for an elliptic finite element problem, Mathematics of Computation 74 (2005) 1–23.

[10] J. Xu, L. Zikatanov, A monotone finite element scheme for convection-diffusion equations, Math. Comput. 68 (1999) 1429–1446.

[11] E. Burman, A. Ern, Discrete maximum principle for Galerkin approximations of the Laplace operator on arbitrary meshes, Comptes Rendus Mathematique 338 (2004) 641 – 646.

[12] E. Burman, A. Ern, Stabilized Galerkin approximation of convection-diffusion-reaction equations: discrete maximum principle and convergence, Mathematics of Computation 74 (2005) 1637–1652.

[13] E. Bertolazzi, G. Manzini, A second-order maximum principle preserving finite volume method for steady convection-diffusion problems, SIAM Journal on Numerical Analysis 43 (2005) 2172–2199.

[14] K. Lipnikov, M. Shashkov, D. Svyatskiy, Y. Vassilevski, Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes, Journal of Computational Physics 227 (2007) 492 – 512.

[15] X. Zhang, C.-W. Shu, Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms, Journal of Computational Physics In Press, Corrected Proof (2010) –.

[16] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, J. Comput. Phys. 229 (2010) 3091–3120.

[17] X. Zhang, C.-W. Shu, On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes, Journal of Computational Physics 229 (2010) 8918 – 8934.

[18] Y. X. X. Zhang, C.-W. Shu, Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes, Technical Report 2010-18, Scientific Computing Group, Brown University, Providence, RI, USA, 2010.

[19] M. Berger, S. M. Murman, M. J. Aftosmis, Analysis of slope limiters on irregular grids, in: Proceedings of the 43rd AIAA Aerospace Sciences Meeting, AIAA2005-0490, AIAA, Reno, NV.

[20] C. Hirt, A. Amsden, J. Cook, An arbitrary Lagrangian-Eulerian computing method for all flow speeds, Journal of Computational Physics 14 (1974) 227–253.

[21] P. Bochev, D. Ridzal, An optimization-based approach for the design of PDE solution algorithms, SIAM Journal on Numerical Analysis 47 (2009) 3938–3955.

[22] P. Bochev, D. Ridzal, Additive operator decomposition and optimization–based reconnection with applications, in: I. Lirkov, S. Margenov, J. Wasniewski (Eds.), Proceedings of LSSC 2009, volume 5910 of *Springer Lecture Notes in Computer Science*.

[23] P. Bochev, D. Ridzal, D. Young, Optimization–based modeling with applications to transport. Part 1. Abstract formulation., in: I. Lirkov, S. Margenov, J. Wasniewski (Eds.), Proceedings of LSSC 2011, Springer Lecture Notes in Computer Science.

[24] P. Bochev, D. Ridzal, G. Scovazi, M. Shashkov, Constrained optimization based data transfer – a new perspective on flux correction, in: D. Kuzmin, R. Lohner, S. Turek (Eds.), Flux-Corrected Transport. Principles, Algorithms and Applications, Springer Verlag, Berlin, Heidelberg, 2nd edition, 2012, pp. 345–398.

[25] L. G. Margolin, M. Shashkov, Second-order sign-preserving conservative interpolation (remapping) on general grids, Journal of Computational Physics 184 (2003) 266 – 298.

[26] M. Kucharik, M. Shashkov, B. Wendroff, An efficient linearity-and-bound-preserving remapping method, Journal of Computational Physics 188 (2003) 462 – 471.

[27] J. K. Dukowicz, J. R. Baumgardner, Incremental remapping as a transport/advection algorithm, Journal of Computational Physics 160 (2000) 318 – 335.

[28] W. H. Lipscomb, E. C. Hunke, Modeling sea ice transport using incremental remapping, Monthly Weather Review 132 (2004) 1341–1354.

[29] P. H. Lauritzen, R. D. Nair, P. A. Ullrich, A conservative semi-lagrangian multi-tracer transport scheme (cslam) on the cubed-sphere grid, Journal of Computational Physics 229 (2010) 1401 – 1424.

[30] R. Liska, M. Shashkov, P. Váchal, B. Wendroff, Optimization-based synchronized flux-corrected conservative interpolation (remapping) of mass and momentum for arbitrary Lagrangian-Eulerian methods, J. Comput. Phys. 229 (2010) 1467–1497.

[31] R. Liska, M. Shashkov, Enforcing the discrete maximum principle for linear finite element solutions of second-order elliptic problems., Commun. Comput. Phys. 3 (2008) 852–877.

[32] J. A. Evans, T. J. Hughes, G. Sangalli, Enforcement of constraints and maximum principles in the variational multiscale method, Computer Methods in Applied Mechanics and Engineering 199 (2009) 61 – 76.

[33] K. B. Nakshatrala, A. J. Valocchi, Non-negative mixed finite element formulations for a tensorial diffusion equation, J. Comput. Phys. 228 (2009) 6726–6752.

[34] O. Angelini, C. Chavant, E. Chénier, R. Eymard, A finite volume scheme for diffusion problems on general meshes applying monotony constraints, SIAM Journal on Numerical Analysis 47 (2010) 4193–4213.

[35] M. J. Berger, P. Colella, Local adaptive mesh refinement for shock hydrodynamics, J. Comput. Phys. 82 (1989) 64–84.

[36] P. Ciarlet, The Finite Element Method for Elliptic Problems, SIAM Classics in Applied Mathematics, SIAM, Philadelphia, 2002.

[37] J. Hyman, M. Shashkov, Natural discretizations for the divergence, gradient, and curl on logically rectangular grids, Computers and Mathematics with Applications 33 (1997) 81 – 104.

[38] P. Bochev, M. Hyman, Principles of mimetic discretizations, in: D. N. Arnold, P. Bochev, R. Lehoucq, R. Nicolaides, M. Shashkov (Eds.), Compatible Discretizations, Proceedings of IMA Hot Topics Workshop on Compatible Discretizations, volume IMA 142, Springer Verlag, 2006, pp. 89–120.

[39] J. Hyman, M. Shashkov, The orthogonal decomposition theorems for mimetic finite difference methods, SIAM Journal on Numerical Analysis 36 (1999) 788–818.

[40] P. L. Roe, Characteristic-based schemes for the euler equations, Annual Review of Fluid Mechanics 18 (1986) 337–365.

[41] D. C. Barth, Timothy J.; Jespersen, The design and application of upwind schemes on unstructured meshes, in: Proceedings of the 27th AIAA Aerospace Sciences Meeting, AIAA89-0366, AIAA, Reno, NV.

[46] R. J. LeVeque, High-resolution conservative algorithms for advection in incompressible flow, SIAM Journal on Numerical Analysis 33 (1996) 627–665.

[42] P. W. Jones, First- and second-order conservative remapping schemes for grids in spherical coordinates, Monthly Weather Review 127 (1999) 2204–2210.

[43] W. H. Lipscomb, T. D. Ringler, An incremental remapping transport scheme on a spherical geodesic grid, Monthly Weather Review 133 (2005) 2335–2350.

[44] D. Ridzal, P. Bochev, J. Young, K. Peterson, Optimization–based modeling with applications to transport. Part 3. Implementation and computational studies., in: I. Lirkov, S. Margenov, J. Wasniewski (Eds.), Proceedings of LSSC 2011, Springer Lecture Notes in Computer Science.

[45] J. Young, D. Ridzal, P. Bochev, Optimization–based modeling with applications to transport. Part 2. Optimization

algorithm., in: I. Lirkov, S. Margenov, J. Wasniewski (Eds.), Proceedings of LSSC 2011, Springer Lecture Notes in Computer Science.

[47] R. D. Nair, P. H. Lauritzen, A class of deformational flow test cases for linear transport problems on the sphere, Journal of Computational Physics 229 (2010) 8868–8887.